



# **The function of respiratory complex I in plants and in human disease**

**Andrew Ewan Maclean**

A thesis submitted to the University of East Anglia for the degree of  
Doctor of Philosophy

John Innes Centre  
Norwich  
September 2017

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived there from must be in accordance with current UK Copyright Law. In addition, any quotation or extract must include full attribution

## Abstract

Complex I is the largest complex in the mitochondrial respiratory chain. Defects in complex I are a major cause of mitochondrial disease in humans. Mutations in the assembly factor *NUBPL* have been implicated in causing complex I deficiency. To assign pathogenicity to patient *NUBPL* variants, I used a yeast model, *Yarrowia lipolytica*, and recreated the corresponding amino acid changes in the Ind1 homolog. Using a combination of BN-PAGE, Western blotting and enzymatic analysis I was able to assign pathogenicity to four of the six variants as well as furthering our understanding of the role of Ind1 in complex I assembly.

Complex I has been lost in the course of evolution in several unicellular eukaryotes, but never in multicellular eukaryotes. Recently, two studies found that the mitochondrial genes encoding complex I subunits were lacking in the genus *Viscum*. To investigate if complex I has been lost, I isolated mitochondria from European Mistletoe, *Viscum album*. My results from BN-PAGE and proteomic analysis indicate that complex I has been lost.

Complex I requires FeS clusters, which are delivered by the mitochondrial ISC pathway. To better understand this process in plants, I characterised the role of FeS carrier proteins NFU4, NFU5 and GRXS15 in *Arabidopsis thaliana*. *NFU4* and *NFU5* were found to be genetically redundant but when combined as a double mutant were embryo lethal. This suggest that NFU4 and NFU5 play an important role in FeS assembly. Mutants in *GRXS15* had a severe growth phenotype, but normal levels of respiratory complexes, suggesting GRXS15 plays a secondary role in FeS cluster assembly.

Understanding complex I will be important in the future for helping to treat human mitochondrial disorders. In addition, studying complex I in plants, including in non-model organisms, helps further our understanding of its function and evolution.

## Acknowledgements

First of all, I would like to express my gratitude to my supervisor, Janneke Balk. I would like to thank her for her guidance and advice throughout my PhD, as well as always having an open door.

I would also like to thank Alison Smith, my secondary supervisor, for her valuable insight and suggestions, as well as Nick Brewin and Steph Bornemann for their running of the rotation PhD program.

Much of the work here has been greatly enriched as the result of collaboration. I would like to thank our collaborators on the “NUBPL project”, Virginia Kimonis, Peggy Eis and Eli Hatchwell for their invaluable input. Many thanks to Etienne Meyer, with whom I worked closely on the “Mistletoe project”, for his advice, expertise and hospitality. I would also like to thank Nicolas Rouhier and Jonathan Pryzbyla-Toscano for their work on the “NFU project”.

For funding, I would like to thank the John Innes Foundation.

I also want to thank the entire Balk group, past and present, as well as the department of biological chemistry. I'd especially like to thank Emma, Inga, James, Jenny, Jorge, Kat and Rob for making lab 108 the best place to be.

I would like to thank the other students at the JIC and TSL for sharing the journey. I would especially like to thank the former residents of “The Marble Factory”, as well as my fellow rotation students, Jemima, Jie, Marc and Nuno.

Special thanks to Jamie, who always knows how to make me laugh.

Last, but not least, I would like to thank my mum, my sister, my dad, my grandparents and Inga for their love and support over the years. I wouldn't be where I am today without them.

# Table of contents

Abstract .....	ii
Acknowledgements.....	iii
Table of contents .....	iv
Table of figures.....	x
Table of tables .....	xii
Abbreviations.....	xiv
1 Function of respiratory complex I in plants and in human disease .....	1
1.1 Mitochondrial research .....	1
1.2 The respiratory chain and oxidative phosphorylation .....	1
1.3 Complex I .....	3
1.3.1 Structure and function of complex I.....	3
1.3.2 Complex I assembly .....	9
1.4 Disease .....	15
1.4.1 Mitochondrial disorders.....	15
1.4.2 Respiratory chain defects .....	16
1.5 Evolution.....	17
1.5.1 Mitochondrial evolution and mitochondrion-related organelles.....	17
1.5.2 Evolution of complex I.....	19
1.5.3 Loss of complex I.....	20
1.6 Cofactors .....	20
1.7 Iron-sulfur clusters .....	21
1.7.1 Iron-sulfur cluster function .....	21
1.7.2 Iron-sulfur cluster assembly .....	21
1.7.3 The mitochondrial ISC pathway .....	23
1.7.4 Iron-sulfur assembly in plants .....	25

1.7.5	Iron-sulfur assembly and human disease.....	25
1.8	Objectives.....	26
2	Materials and methods .....	27
2.1	Bioinformatics.....	27
2.1.1	Databases .....	27
2.1.2	Alignments.....	27
2.1.3	Homology modelling .....	27
2.1.4	Illustration of complex I structure .....	27
2.1.5	Prediction of functional effects of human nsSNPs.....	27
2.2	Chemicals.....	27
2.3	Organisms .....	28
2.3.1	<i>Arabidopsis thaliana</i> .....	28
2.3.2	<i>Yarrowia lipolytica</i> .....	28
2.3.3	<i>Escherichia coli</i> .....	29
2.4	Molecular techniques: Polymerase Chain Reaction .....	29
2.4.1	DNA extraction from <i>Arabidopsis thaliana</i> .....	29
2.4.2	DNA extraction from <i>Yarrowia lipolytica</i> .....	29
2.4.3	RNA extraction from <i>Arabidopsis thaliana</i> .....	30
2.4.4	RNA extraction from <i>Yarrowia lipolytica</i> .....	30
2.4.5	cDNA synthesis and reverse transcription .....	30
2.4.6	Polymerase chain reaction .....	30
2.4.7	DNA extraction from agarose gel.....	30
2.4.8	Agarose gel electrophoresis .....	31
2.4.9	Oligonucleotides.....	31
2.5	Molecular techniques: Plasmid construction .....	34
2.5.1	Restriction digest.....	34
2.5.2	Plasmid isolation.....	34
2.5.3	DNA sequencing.....	34

2.5.4	Ligation.....	34
2.5.5	Mutagenesis .....	34
2.5.6	Protein expression vectors .....	34
2.5.7	Plasmid created and used in this study.....	35
2.6	Genetic transformation .....	35
2.6.1	<i>Escherichia coli</i> transformation .....	35
2.6.2	<i>Yarrowia lipolytica</i> transformation .....	36
2.6.3	Antibiotics .....	36
2.7	<i>Arabidopsis thaliana</i> growth and manipulation .....	36
2.7.1	Growth on MS plates .....	36
2.7.2	Growth on soil .....	36
2.7.3	Growth in hydroponic culture .....	37
2.7.4	Callus cell lines.....	37
2.7.5	Cross pollination .....	37
2.8	<i>Yarrowia lipolytica</i> growth .....	37
2.8.1	Agar plates .....	37
2.8.2	Liquid culture .....	37
2.9	Mitochondrial purification .....	38
2.9.1	Mitochondrial purification from <i>Arabidopsis thaliana</i> leaves or hydroponic culture .....	38
2.9.2	Mitochondrial purification from <i>Arabidopsis thaliana</i> callus .....	39
2.9.3	<i>Yarrowia lipolytica</i> mitochondrial membranes .....	39
2.10	SDS-PAGE and Western blotting.....	40
2.10.1	Total protein isolation <i>Arabidopsis thaliana</i> and <i>Viscum album</i> .....	40
2.10.2	Protein determination.....	40
2.10.3	SDS-PAGE .....	40
2.10.4	Immunoblotting and ECL detection.....	41
2.10.5	Ponceau and InstantBlue™ stain.....	41

2.10.6	Antibodies.....	42
2.11	Blue Native PAGE .....	42
2.11.1	<i>Arabidopsis thaliana</i> and <i>Viscum album</i> .....	42
2.11.2	<i>Yarrowia lipolytica</i> .....	43
2.11.3	Transfer to PVDF membrane and immunolabelling .....	43
2.11.4	Coomassie stain .....	43
2.11.5	Complex I, II and IV activity stain .....	43
2.11.6	Complex III detection.....	44
2.12	Spectrophotometric enzyme assays .....	44
2.12.1	Complex I NADH:HAR oxidoreductase assays .....	44
2.12.2	Complex I dNADH:DBQ oxidoreductase assay.....	44
2.13	Microscopy .....	46
2.13.1	Light microscopy.....	46
2.13.2	Differential Interference Contrast .....	46
2.14	Proteomics .....	46
2.14.1	Sample preparation and mass spectrometry.....	46
2.14.2	Analysis.....	47
2.15	Protein expression.....	47
3	Using <i>Yarrowia lipolytica</i> as a model to understand pathogenic mutations in <i>NUBPL</i> associated with complex I deficiency .....	48
3.1	Introduction.....	48
3.2	Results .....	51
3.2.1	Selection of <i>NUBPL</i> mutations for study .....	51
3.2.2	Bioinformatic analysis of <i>NUBPL</i> variants .....	51
3.2.3	Protein stability of Ind1 variants .....	57
3.2.4	Variants L102P, D103Y and L191F are unable to support cWT levels of complex I assembly .....	58
3.2.5	The levels of NUBM and NUCM, two subunits of Complex I, are not significantly altered .....	61

3.2.6	<i>ind1</i> $\Delta$ accumulates the Q module assembly intermediate .....	63
3.2.7	All Ind1 variants accumulate the Q module of complex I.....	65
3.2.8	Ind1 variants are sensitive to cold.....	68
3.2.9	Complex I is unstable at cold temperatures for Ind1 variant G285C.....	68
3.2.10	Semi-conserved NUBPL variant V182A does not affect complex I assembly .....	70
3.2.11	Studies on recombinant Ind1 show L102P is insoluble .....	72
3.3	Discussion .....	74
4	Life without complex I in a parasitic plant.....	77
4.1	Introduction.....	77
4.2	Results .....	79
4.2.1	Isolation of mitochondria from <i>Viscum album</i> .....	79
4.2.2	BN-PAGE and in gel-activity assay indicate <i>Viscum album</i> lacks complex I 83	
4.2.3	BN-PAGE and in gel-activity assays suggest <i>Viscum album</i> retains complex II, III and IV .....	83
4.2.4	Proteomic analysis indicate that complex I is absent, but the rest of the respiratory chain is intact .....	86
4.2.5	Analysis and comparison of <i>Arabidopsis thaliana</i> complex I mutants.....	90
4.2.6	Further characterisation of <i>Arabidopsis thaliana indh</i> mutant.....	92
4.3	Discussion .....	95
5	Functional characterization of FeS assembly proteins in <i>Arabidopsis thaliana</i> mitochondria .....	98
5.1	Introduction.....	98
5.2	Results .....	103
5.2.1	Characterisation of the <i>isu1-1</i> mutant .....	103
5.2.2	NFU4 and NFU5 are localised to the mitochondrial matrix.....	105
5.2.3	Genetic analysis of <i>NFU4</i> and <i>NFU5</i> mutants .....	107



5.2.4	<i>NFU4</i> and <i>NFU5</i> mutants display no growth phenotype suggesting functional redundancy.....	111
5.2.5	<i>NFU4</i> and <i>NFU5</i> are essential for embryo development.....	111
5.2.6	Single mutants of <i>NFU4</i> and <i>NFU5</i> have no noticeable FeS assembly phenotype.....	114
5.2.7	Hemizygous <i>NFU</i> mutant has a decrease in H-protein .....	117
5.2.8	<i>GRXS15</i> mutants do not have a strong FeS assembly phenotype.....	120
5.3	Discussion .....	123
6	Discussion .....	126
6.1	The function of Ind1 in complex I assembly .....	126
6.1.1	The role of Ind1 is evolutionarily conserved but the molecular function is unclear	126
6.1.2	A direct role in complex I assembly.....	127
6.1.3	A role for INDH in mitochondrial translation .....	128
6.2	FeS provision for the respiratory complexes .....	129
6.2.1	Defects in the ISC pathway impact on respiratory complex assembly..	129
6.2.2	LYR motif proteins involved with FeS delivery to respiratory complexes .....	131
6.3	Distinguishing between two models of FeS cluster assembly in the mitochondria .....	132
6.4	<i>Yarrowia lipolytica</i> as a model for complex I assembly .....	133
7	Bibliography.....	135
8	Appendix .....	155
	Proteomic data from Chapter 4 .....	155

## Table of figures

Figure 1.1: Function of complex I.....	4
Figure 1.2: Structure of complex I from <i>Bos taurus</i> .....	5
Figure 1.3: Complex I assembly in <i>Homo sapiens</i> .....	11
Figure 1.4: Comparison of complex I assembly in <i>Homo sapiens</i> and <i>Yarrowia lipolytica</i> .....	14
Figure 1.5: FeS assembly pathways in the eukaryotic cell .....	22
Figure 2.1: Mitochondrial purification using a Percoll gradient.....	38
Figure 2.2: Histochemical stains of the mitochondrial respiratory chain .....	45
Figure 3.1: Alignment of NUBPL and Ind1 .....	54
Figure 3.2: Homology model of Ind1 and amino acids investigated in this study .....	54
Figure 3.3: Expression of variant Ind1 proteins in <i>Yarrowia lipolytica</i> .....	56
Figure 3.4: Complex I levels and oxidoreductase activity in Ind1 variants .....	59
Figure 3.5: Complex I subunits in complex I mutants and Ind1 variants .....	62
Figure 3.6: Complex I assembly intermediates in Ind1 and Complex I mutants.....	64
Figure 3.7: Complex I assembly intermediates in Ind1 variants.....	66
Figure 3.8: <i>Yarrowia lipolytica</i> growth at low temperature .....	67
Figure 3.9: Cold temperatures lead to complex I instability in G285C variant.....	69
Figure 3.10: Complex I levels and growth in M180 Ind1 variants .....	71
Figure 3.11: Ind1 and Ind1 variant protein expressed in <i>Escherichia coli</i> .....	71
Figure 4.1: Purification of mitochondria from <i>Viscum album</i> .....	81
Figure 4.2: Western blots of <i>Viscum album</i> mitochondrial purifications .....	82
Figure 4.3: Complex I activity stain is absent in <i>Viscum album</i> .....	84
Figure 4.4: Complex II, III and IV in-gel activity in <i>Viscum album</i> .....	85
Figure 4.5: Proteins in <i>Viscum album</i> mitochondria detected by mass spectrometry ..	88
Figure 4.6: Comparison of complex I in <i>Arabidopsis thaliana</i> mutants .....	91
Figure 4.7: Further characterisation of <i>Arabidopsis thaliana indh</i> mutant .....	93
Figure 5.1: Mitochondrial ISC pathway in <i>Arabidopsis thaliana</i> .....	102

Figure 5.2: Analysis of <i>isu1-1</i> mutant.....	104
Figure 5.3: NFU4 and NFU5 are localised in the mitochondrial matrix .....	106
Figure 5.4: Genetic analysis of <i>Arabidopsis thaliana</i> mutants in <i>NFU4</i> and <i>NFU5</i> ....	108
Figure 5.5: Growth analysis <i>NFU4</i> and <i>NFU5</i> T-DNA insertion lines.....	109
Figure 5.6: Isolation of <i>nfu5-1</i> line.....	110
Figure 5.7: Investigation of <i>NFU4 NFU5</i> double mutant.....	113
Figure 5.8: Western blot and BN-PAGE analysis of <i>NFU4</i> and <i>NFU5</i> single mutants	116
Figure 5.9: Creation of the hemizygous <i>NFU</i> mutant .....	118
Figure 5.10: BN-PAGE and Western blot analysis of hemizygous and <i>nfu4-2 nfu5-2</i> double mutants .....	119
Figure 5.11: Analysis of <i>grxs15-K83A</i> and <i>amiRNA grxs15</i> mutants.....	122

## Table of tables

Table 1.1: Subunit composition of the respiratory chain in <i>Homo sapiens</i> .....	4
Table 1.2: Core subunits of complex I.....	6
Table 1.3: Complex I subunits in <i>Homo sapiens</i> , <i>Yarrowia lipolytica</i> and <i>Arabidopsis thaliana</i> .....	6
Table 1.4: Complex I assembly factors in <i>Homo sapiens</i> .....	12
Table 2.1: List of <i>Arabidopsis thaliana</i> lines .....	28
Table 2.2: List of <i>Yarrowia lipolytica</i> strains .....	28
Table 2.3: List of <i>Escherichia coli</i> strains .....	29
Table 2.4: Primers used for RT-PCR .....	31
Table 2.5: Primers used for genotyping .....	32
Table 2.6: Primers used for mutagenesis.....	33
Table 2.7: Primers used for cloning and sequencing.....	33
Table 2.8: List of plasmids .....	35
Table 2.9: List of antibiotics .....	36
Table 2.10: SDS-PAGE gel composition.....	40
Table 2.11: ECL reagents .....	41
Table 2.12: List of antibodies .....	42
Table 2.13: BN-PAGE in gel activity stains .....	44
Table 3.1: Overview of NUBPL mutations associated with complex I deficiency .....	52
Table 3.2: Protein variants in <i>Homo sapiens</i> NUBPL and <i>Yarrowia lipolytica</i> Ind1 .....	53
Table 3.3: PolyPhen-2 predictions of effect of NUBPL variants.....	53
Table 3.4: Allele frequency of mutations leading to NUBPL variants from ExAC .....	53
Table 3.5: Complex I activity assays for Ind1 variants.....	60
Table 3.6: Complex I activity in patients.....	73
Table 3.7: Complex I activity assays for Ind1 variants G80S and K81Q.....	73
Table 4.1: <i>Viscum album</i> sample collection .....	81
Table 4.2: Proteins identified by comparison to <i>Viscum album</i> mtDNA .....	89

Table 4.3: Proteins identified by comparison to <i>Arabidopsis thaliana</i> database .....	89
Table 5.1: Mitochondrial ISC pathway in <i>Arabidopsis thaliana</i> .....	102
Table 5.2: Summary of mutant alleles .....	106
Table 5.3: Embryo abortion counts of <i>NFU</i> mutants.....	113
Table 8.1: mtDNA encoded mitochondrial proteins from <i>Viscum album</i> Percoll-purified mitochondria .....	155
Table 8.2: mtDNA encoded mitochondrial proteins from <i>Viscum album</i> crude mitochondria .....	155
Table 8.3: Mitochondrial proteins identified based on the <i>Arabidopsis thaliana</i> database in <i>Viscum album</i> Percoll purified mitochondria.....	156
Table 8.4: Mitochondrial proteins identified based on the <i>Arabidopsis thaliana</i> database in <i>Viscum album</i> crude mitochondria .....	156

## Abbreviations

ABC	ATP-binding cassette
APS	Ammonium persulphate
ATP	Adenosine triphosphate
ATPase	Adenosine triphosphate hydrolase
Bp	Base pairs
BN	Blue native
CII	Complex II
CIII	Complex III
CIV	Complex IV
CBB	Coomassie brilliant blue
cDNA	Complementary DNA
CIA	Cytosolic iron-sulfur protein assembly
CoA	Coenzyme A
Col-0	Columbia-0
cWT	Complemented wild type
DAB	3,3'-Diaminobenzidine
DBQ	n-decylubiquinone
dH <sub>2</sub> O	Distilled water
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
dNADH	Deamino NADH
dNTPs	deoxynucleotides

DTT	Dithiothreitol
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EPR	Electron paramagnetic resonance
e.v	Empty vector
F2	Second filial generation
FAD	Flavin adenine dinucleotide
FeS	Iron-sulfur
FMN	Flavin mononucleotide
Gb	Giga base pairs
GDC	Glycine decarboxylase
HAR	Hexaammineruthenium(III) chloride
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
ISC	Iron-sulfur cluster assembly
LB	Lysogeny broth/ Luria-Bertani
Ler	Landsberg erecta
MES	2-(N-morpholino)ethanesulfonic acid
MOPS	3-(N-morpholino)propanesulfonic acid
mRNA	Messenger RNA
MS	Murashige and skoog
mt	Mitochondrial
MTS	Mitochondria targeting sequence
NAD <sup>+</sup>	Nicotinamide adenine dinucleotide (oxidised)
NADH	Nicotinamide adenine dinucleotide (reduced)

NADPH	Nicotinamide adenine dinucleotide phosphate (reduced)
NBT	Nitro-blue tetrazolium
nsSNPs	Non-synonymous single nucleotide polymorphisms
NTPase	Nucleoside triphosphate hydrolase
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDH	Pyruvate dehydrogenase
PMS	Phenazine methosulfate
PMSF	Phenylmethanesulfonyl fluoride
qRT-PCR	Quantitative reverse transcriptase polymerase chain reaction
RNA	Ribonucleic acid
RT-PCR	Reverse transcriptase polymerase chain reaction
S <sup>0</sup>	Persulfide
s.d.	Standard deviation
SDS	Sodium dodecyl sulphate
s.e.m	Standard error of the mean
SOC	Super optimal broth with catabolite repression
SUF	Sulfur mobilisation
TAE	Tris-acetic acid-EDTA
TBS	Tris-buffered saline
TBS-TM	Tris-buffered saline with milk and tween
TE	Tris-EDTA
TEMED	Tetramethylethylenediamine



T-DNA	Transfer DNA
Tris	Tris(hydroxymethyl)aminomethane
tRNAs	Transfer RNA
UTR	Untranslated region
UV	Ultraviolet
v/v	Volume/volume
WT	Wild type
w/v	Weight/volume

# **1 Function of respiratory complex I in plants and in human disease**

## **1.1 Mitochondrial research**

Mitochondria have played an important role in the development of modern cell biology and biochemistry. Inside mitochondria resides the biochemical machinery needed for respiration, some of life's most important and central chemical reactions. This has led mitochondria to be frequently referred to as "the powerhouses of the cell". Unravelling the mechanisms and intricate machinery involved in respiration was one of the great achievements of 20<sup>th</sup> century science. It was a central subject and many of the great minds of the age spent careers teasing out and unravelling the details. Indeed, numerous Nobel prizes were awarded for subjects relating to respiration and mitochondria. These include Otto Warburg in 1931, for his work on respiratory enzymes, Hans Krebs in 1953 for work on the Krebs cycle, Peter Mitchell in 1978 for developing the chemiosmotic theory and Paul Boyer and John Walker in 1997 for their work on ATP synthase. Mitochondrial research in the second half of the century did not capture the imaginations quite as before, despite numerous important and ground-breaking discoveries. However mitochondrial research has undergone a renaissance since the beginning of the century, and once again is a subject of central interest.

## **1.2 The respiratory chain and oxidative phosphorylation**

Mitochondria are semi-autonomous organelles which have two membranes, the inner of which forms the characteristic cristae and encloses the mitochondrial matrix. Mitochondria are the home of oxidative phosphorylation, which is involved in cellular ATP production. ATP synthesis is a prime example of the chemiosmotic theory, put forward by Peter Mitchell in 1961, and is widely considered to be one of the great scientific breakthroughs of the 20<sup>th</sup> century. Stated simply, the process of oxidative phosphorylation is as follows: pyruvate, generated from glycolysis, undergoes oxidative decarboxylation, to form acetyl-CoA and CO<sub>2</sub>, which in turn is sequentially oxidised and decarboxylised by the Krebs, or tricarboxylic (TCA), cycle, which has been described as "an asset-stripping merry-go round of linked reactions" (Lane 2005). Electrons harvested from the Krebs cycle electrons are then fed into the respiratory chain (see next paragraph). The electrons are then passed along the chain in a succession of linked redox reactions (i.e. each electron carrier is first reduced, and then oxidised by the next link in the chain). The chain ends with the reduction of oxygen. These successive redox

reactions are coupled to the transport of protons across the inner mitochondrial membrane, from the matrix to the inter-membrane space. This proton gradient produces the proton-motive-force (see next paragraph) which is then used by the enzyme complex ATP synthase to generate ATP, the common currency of energy in the cell, which allows energy to be stored and then used at a later time, often in a different part of the cell. The action of ATP synthase, using the proton-motive force to generate ATP, is often metaphorically compared to a hydro-electric dam, where potential energy is converted into electrical energy, which can then be put to use.

A basic summary of the process belies the underlying complexity of this remarkable piece of biochemistry. Looking at intricate mechanisms of the respiratory chain allows for a greater appreciation and understanding. The respiratory chain is composed of five macro-molecular complexes (the first four complexes are sometimes called the electron transport chain) named: complex I (also known as NADH dehydrogenase and NADH:ubiquinone oxidoreductase), complex II (succinate dehydrogenase or succinate:ubiquinone oxidoreductase), complex III (cytochrome *bc*<sub>1</sub> complex or ubiquinone:cytochrome *c* oxidoreductase), complex IV (cytochrome *c* oxidase) and complex V (ATP synthase). These are composed of approximately 90 proteins, the number varies species to species, encoded by either the mitochondrial or nuclear genome (Table 1.1). The entry point for electrons is through complex I or complex II. Electrons that enter complex I are carried from the Krebs cycle by the reduced form of the electron acceptor nicotinamide adenine dinucleotide (NADH), which is then oxidised, donating two electrons to the Flavin mononucleotide (FMN) cofactor of complex I. These electrons are then passed along an iron-sulfur (FeS) cluster chain, before reducing ubiquinone to ubiquinol. Also, electrons can enter complex II via succinate, which is oxidised, donating its electrons to a flavin adenine dinucleotide (FAD) cofactor. These again are passed along an FeS cluster chain, and reduce ubiquinone to ubiquinol. Additionally, ubiquinone can be reduced by another pathway, fatty acid  $\beta$ -oxidation. Ubiquinone is a small lipid-soluble benzoquinone and so can diffuse through the lipid bilayer to the next electron carrier, complex III. Complex III transfers electrons from ubiquinol, via heme cofactors and a FeS cluster, to another mobile electron carrier, cytochrome *c* in the intermembrane space. It accepts an electron using its heme cofactor and transfers it to the final complex in the electron transport chain, complex IV. Complex IV transfers electrons from cytochrome *c* via heme and copper groups, to molecular oxygen, O<sub>2</sub>, the final electron acceptor, thereby reducing it to H<sub>2</sub>O. The transfer of electrons is coupled to proton pumping by complexes I, III and IV. This creates a proton-motive force ( $\Delta p$ ), also called the chemiosmotic, or electrochemical, gradient. The

proton-motive force is made from the energy stored in electrical charge separation and the pH difference across the inner mitochondrial membrane. The electrochemical gradient in respiring mitochondria is typically between 150 and 200 mV. The proton-motive force is used by the fifth respiratory complex, ATP synthase, to generate ATP through rotational catalysis. Theoretically the complete oxidation of one pyruvate molecule can lead to the generation of 12 ATP molecules.

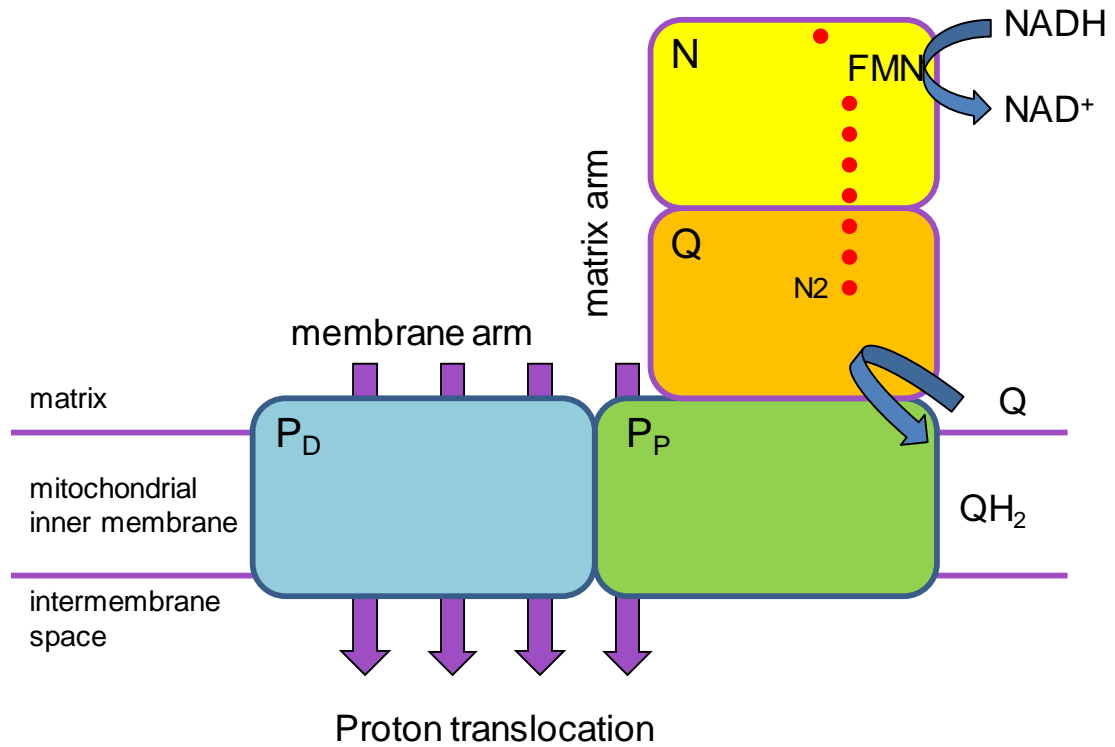
The details of oxidative phosphorylation have been steadily unpicked over the last 50 years. Now this knowledge is found in every textbook and undergraduate course. However, the respiratory chain is not a closed area of research. Many new discoveries are being, and are waiting to be, made. For example, advances in structural biology have revealed in detail the mechanism by which these respiratory complexes function. Just in the last year high resolution cryo-electron microscopy structures of the largest complex, complex I, were solved (Zhu et al. 2016; Fiedorczuk et al. 2016). It is this complex which will be the focus of this thesis.

### **1.3 Complex I**

#### **1.3.1 Structure and function of complex I**

Complex I, also known as NADH dehydrogenase and NADH:Ubiquinone oxidoreductase, is the largest of the respiratory complexes. As mentioned, complex I is the major entry point for electrons into the respiratory chain. Complex I oxidises NADH, taking two electrons, which it uses to reduce ubiquinone to ubiquinol. This redox reaction is coupled to the transfer of protons across the mitochondrial inner membrane against the chemiosmotic gradient. It is generally thought that four protons are translocated for every NADH oxidised (i.e. for every two electrons (Hirst 2013; Jones et al. 2017)). This accounts for approximately 40% of the protons pumped across the inner mitochondrial membrane during oxidative phosphorylation (Berrisford et al. 2016).

Complex I is composed of 40+ protein subunits, the exact composition varies from species to species, which are formed into an L-shaped molecule, with one arm extending into the mitochondrial matrix and the other residing in the mitochondrial inner membrane. In humans, there are 44 different subunits. The number 45 is commonly quoted as well. It was previously thought that there were 45 subunits, beginning with the initial discovery

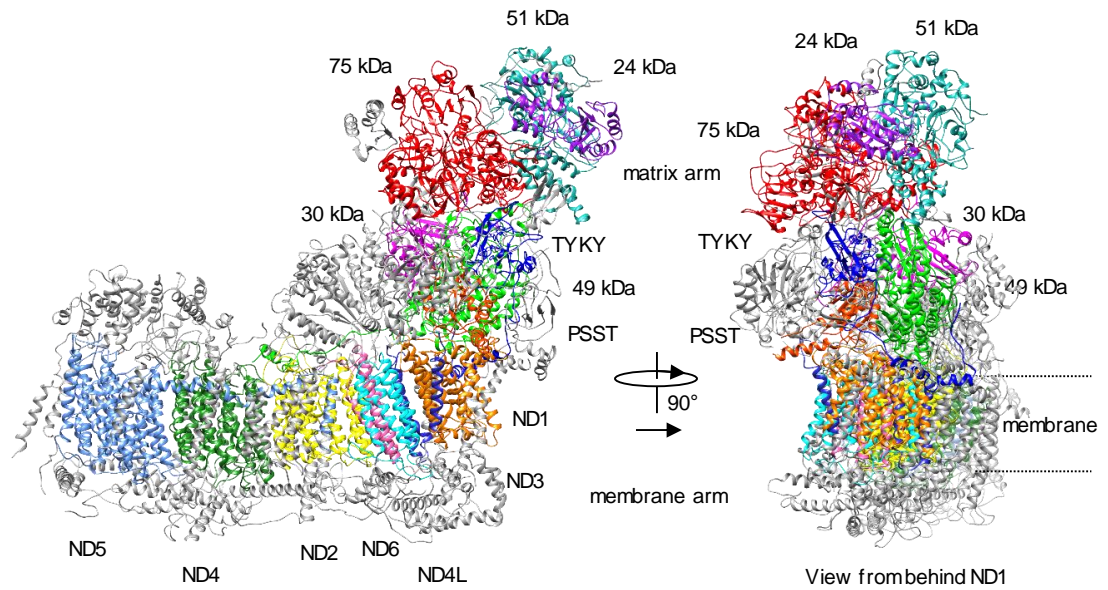


**Figure 1.1: Function of complex I**

Schematic outlining the function of complex I. The N, NADH binding, Q, ubiquinone binding, and P, proton translocation, modules are highlighted. The P module is further divided into the proximal, P<sub>P</sub>, and distal, P<sub>D</sub>, sections. FeS clusters are indicated in red, and proton translocation indicated by the purple arrows.

**Table 1.1: Subunit composition of the respiratory chain in *Homo sapiens***

Complex	Number of subunits	Nuclear encoded	Mitochondrially encoded
I	44	37	7
II	4	4	0
III	11	10	1
IV	13	10	3
V	18	16	2



**Figure 1.2: Structure of complex I from *Bos taurus***

Complete structure of complex I from *Bos taurus* (Protein Data Bank code: 5LDW) at 4.27Å. Core subunits, outlined in table 1.2, are highlighted in colour and labelled. Accessory subunits are in grey. Two orientations are shown, with the view on the right rotated 90° and then viewed from behind the ND1 subunit.

**Table 1.2: Core subunits of complex I**\* Subunits encoded in the mitochondrial genome in *Arabidopsis thaliana*

Genomic origin	Module	<i>Bos taurus</i>	<i>Homo sapiens</i>	<i>Yarrowia lipolytica</i>	<i>Arabidopsis thaliana</i>	Cofactors and $\alpha$ -helices (TMHs)
Nuclear	N	75 kDa	NDUFS1	NUAM	75 kDa	[2Fe-2S], 2 x [4Fe-4S]
		51 kDa	NDUFV1	NUBM	51 kDa	Flavin, [4Fe-4S]
		24 kDa	NDUFV2	NUHM	24 kDa	[2Fe-2S]
	Q	PSST	NDUFS7	NUKM	PSST/ 20 kDa	[4Fe-4S] (cluster N2)
		TYKY	NDUFS8	NUIM	TYKY/ 23 kDa	2 x [4Fe-4S]
		49 kDa	NDUFS2	NUCM	Nad7/ 49 kDa*	
		30 kDa	NDUFS3	NUGM	Nad9/ 30 kDa*	
Mitochondrial	P	ND1	ND1	NU1M	Nad1	8 TMHs
		ND2	ND2	NU2M	Nad2	11-14 TMHs
		ND3	ND3	NU3M	Nad3	3 TMHs
		ND4	ND4	NU4M	Nad4	14 TMHs
		ND5	ND5	NU5M	Nad5	16 TMHs
		ND6	ND6	NU6M	Nad6	5 TMHs
		ND4L	ND4L	NULM	Nad4L	3 TMHs

**Table 1.3: Complex I subunits in *Homo sapiens*, *Yarrowia lipolytica* and *Arabidopsis thaliana***

Based on Angerer et al. 2011, Subrahmanian et al 2016. and Wirth et al. 2016

Species	Subunit number	Nuclear encoded	Mitochondrial encoded
<i>Homo sapiens</i>	44	37	7
<i>Yarrowia lipolytica</i>	42	35	7
<i>Arabidopsis thaliana</i>	49	40	9

that there were 45 subunits in *Bos taurus* complex I (Carroll et al. 2006), but two recent studies in 2012 (Balsa et al. 2012) and 2013 (Pitceathly et al. 2013) suggest that one of these, NDUFA4, is in fact a subunit of complex IV. Later it was found that although there are 44 different subunits, in the structure of *Bos taurus* complex I there are two copies of the SDAP (human NDUFAB1) subunit (Vinothkumar et al. 2014), bringing the total number of individual subunits to 45. The numbers 44 and 45 both continue to be commonly cited, but for the remainder of this thesis 44 will be used. 37 of these subunits are encoded in the nuclear genome and 7 in the mitochondrial genome. This implies coordination between two separate genomes for proper assembly and function of complex I, which adds an extra layer of complexity to its biogenesis. The subunits that are encoded by the mitochondrial genome vary from species to species, but always include ND1, ND4 and ND5 (Brandt 2006). The hydrophobic properties of the mitochondrially encoded subunits has often been given as the reason they are retained in the mitochondrial genome. It was suggested that their hydrophobicity could make their import into the mitochondria difficult, but recently it has been suggested their hydrophobicity could also lead to mis-targeting to the endoplasmic reticulum (Björkholm et al. 2015).

The 40+ subunits of complex I can be divided into two classes: the core, or central, subunits and the accessory, or supernummary, subunits. There are 14 core subunits and these are invariant between species (Table 1.2). These subunits are required for catalysis, and are therefore highly conserved. Indeed, homologs of these subunits exist in bacteria, and they form the minimal functioning form of the enzyme. The core subunits form into two main domains. The seven hydrophilic subunits, which are nuclear encoded, constitute the redox domain, sometimes called the peripheral, or matrix, arm, which extends into the matrix of the mitochondria (Figure 1.1 and 1.2). The redox domain is further divided into the N, NADH binding, i.e. the electron input, and the Q, ubiquinone binding, i.e. the electron output, module (Brandt 2006). It is this domain which contains the cofactors of complex I, which are needed for the redox reactions (see next paragraph).

In humans, the redox domain of complex I contains 9 cofactors, which are used to perform complex I's redox reactions. These are a Flavin mononucleotide (FMN), and 8 FeS clusters (see section 1.7). The FMN participates in NADH oxidation, and passes the electrons to the FeS clusters. The FeS clusters are arranged in a chain, analogous to an electrical wire, along which the electrons are passed. It consists of 1  $[2\text{Fe-2S}]^{2+/1+}$  and 6  $[4\text{Fe-4S}]^{2+/1+}$  clusters, named after their identification by electron paramagnetic resonance (EPR) in the 1970's (Ohnishi 1998). In addition is a  $[2\text{Fe-2S}]^{2+/1+}$  positioned



outside the FeS cluster chain, called N1a, whose function is still uncertain, although it has recently been suggested to confer stability to complex I and aid incorporation of the N1b cluster (Dörner et al. 2017). The final FeS cluster in the chain, called the N2 cluster, donates electrons to ubiquinone, thereby reducing it to ubiquinol. The N2 cluster is attached to the PSST subunit and is positioned  $\sim 12\text{\AA}$  away from ubiquinone, which sits in a cavity in the complex created by subunits 49 kDa and PSST (Wirth et al. 2016). The ubiquinone headgroup interacts with residues H92 and Y141 (human numbering) of subunit NDUFS2 (bovine 49 kDa). However, the exact position and interactions of ubiquinone in complex I remains uncertain and is still being studied. Additionally, the mechanism coupling the redox reactions to proton translocation is currently undefined (Hirst 2013).

The seven hydrophobic subunits, which are mitochondrially encoded, the ND subunits, make up the membrane domain, sometimes called the P, the proton translocation, module, and are embedded in the mitochondrial inner membrane (Figure 1.1 and 1.2). The P module is further divided into the proximal P module ( $P_P$ ), which is connected into the Q module, and the distal P module ( $P_D$ ). Three ND subunits, ND5, ND4 and ND2, are homologous to the Mrp  $\text{Na}^+/\text{H}^+$  antiporter family, and are probably involved in proton pumping (Mathiesen & Hägerhäll 2002; Hunte et al. 2010; Moparthi et al. 2011; Zickermann et al. 2015). Additionally, domains from subunits ND1, ND6 and ND4L form a fourth pathway for proton translocation (Hirst 2013).

The two 'arms' of complex I, the matrix arm and membrane arm, are held together by the interactions of subunits on the interface between them. Hydrogen bonds and salt bridges form between the membrane domain subunit, ND1, and the redox domain subunit 49 kDa, PSST and TYKY. Additionally, a protein loop from the membrane domain subunit ND3 extends up and interacts with the interface of subunits 49 kDa and PSST (Hirst 2013).

In addition to the core subunits are the accessory, or supernumerary, subunits (Kmita & Zickermann 2013). These are mostly absent from the bacterial complex I, and result in an approximate doubling of the mass of the eukaryotic complex (Wirth et al. 2016). In humans, there are 30 of these, but the number varies between species. For example, in the aerobic yeast *Yarrowia lipolytica* there are 28 accessory subunits (Kmita & Zickermann 2013), while the model plant *Arabidopsis thaliana* has 35 (Subrahmanian et al. 2016) (Table 1.3). In eukaryotes, these are all encoded by the nuclear genome. Accessory subunits are defined as those that do not participate directly in catalysis. Some of these may play structural roles, or may help complex I assembly or regulation,

or may protect against reactive oxygen species. The function of accessory subunits is still not well understood. A recent study used gene editing technology to create knock-out mutants of all accessory subunits in human cell lines, and showed that 25, out of 30, are essential for proper formation of complex I (Stroud et al. 2016). This study provides a striking example of the importance of accessory subunits, and highlights the need for further study of their function.

The nomenclature of complex I subunits has long been a complicated issue. There is no unified naming convention, and subunits often have different names. The first described subunits were from *Bos taurus*, and were so called by their apparent size on a SDS-PAGE gel (e.g. 75 kDa subunit). It has become convention to refer to subunits by their homology to the *Bos taurus* subunits. Here, subunits will be mainly referred to by their preferred names in the individual species of study. Table 1.2 can be used to track which subunit is homologous to which.

Great advances have been made in the last decade in elucidating the structure of complex I. Obtaining a precise structure is essential for fully understanding the reaction mechanism of complex I, for example by showing the exact position of key catalytic residues or clearly defining reaction chambers. Structures of complex I have been obtained from diverse phylogenetic groups across the eukaryotes and bacteria, with structures from the bacteria, *Thermus thermophilus* (Baradaran et al. 2013), the fungi *Yarrowia lipolytica* (Zickermann et al. 2015) and mammals, from *Bos taurus* (Zhu et al. 2016) and *Ovis aries* (Fiedorczuk et al. 2016). Great progress has been made, moving from structures of individual domains up to high resolution structures of the entire complex. The similarity of structure across these diverse group once again highlights the conservation of complex I structure and function.

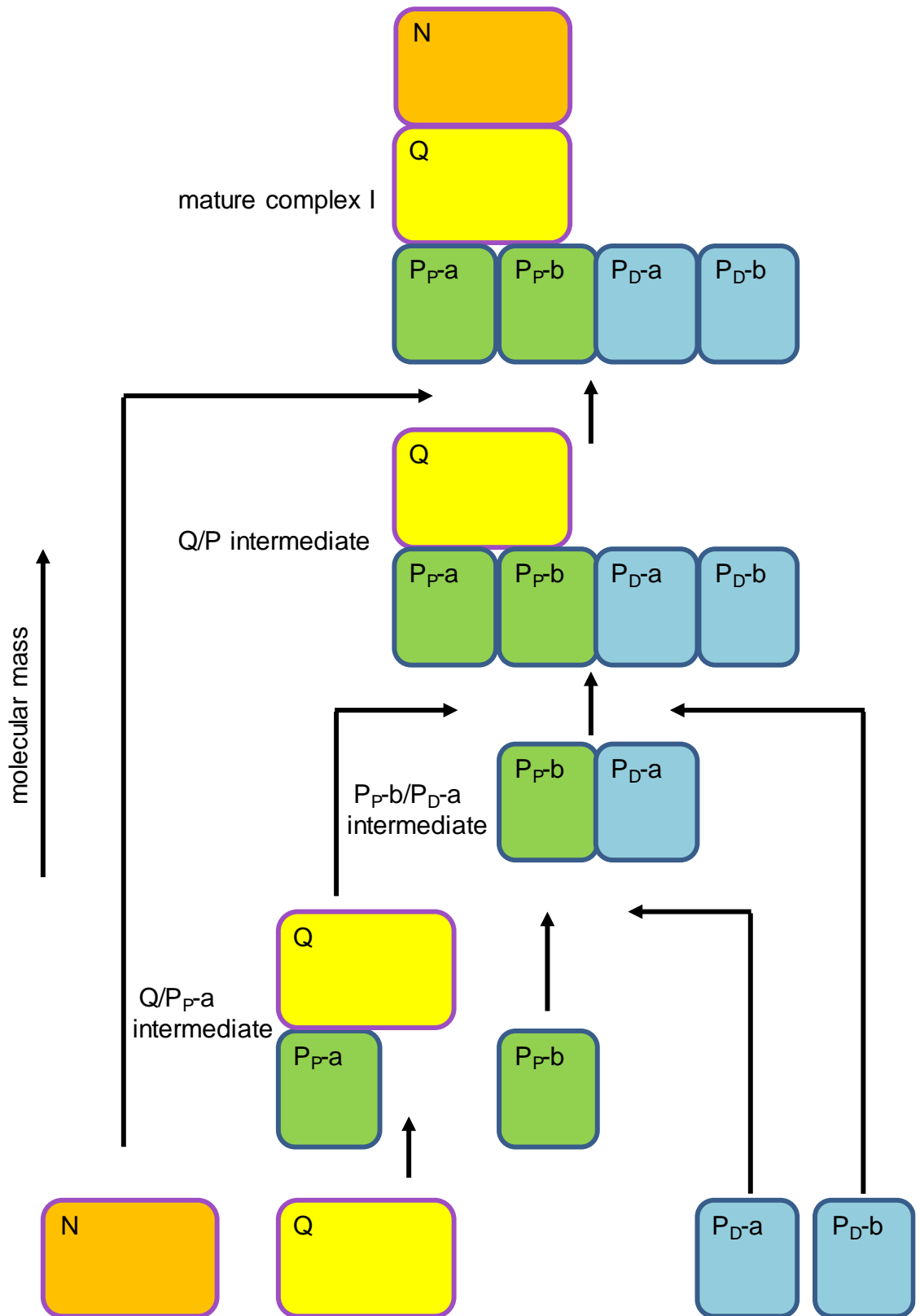
### 1.3.2 Complex I assembly

Given complex I's large size, the origin of subunits from two genomes and the cofactors that need to be properly inserted, complex I assembly is a complicated, multi-stage process that is tightly controlled. This complicated process has been recently reviewed in (Sánchez-Caballero et al. 2016) and (Formosa et al. 2017) and will be outlined below. Note, the exact sequence of events varies between species (Vogel et al. 2007), but the process is best studied in humans and so will be described first.

The seven mitochondrially encoded subunits require the correct transcription and translation of the mitochondrial genome, itself a complex process with many parts. Mitochondrial ribosomes are located in close proximity to the inner mitochondrial

membrane. This allows co-translational insertion into the membrane, with the help of the OXA1 protein (Sánchez-Caballero et al. 2016). The remainder of complex I subunits, as well as the 16 assembly factors, are encoded in the nuclear genome. These are translated in cytosolic ribosomes before transport into the mitochondria. Most, but not all, of these proteins contain N-terminal mitochondrial targeting sequences (MTS) to ensure proper delivery, which are later removed. Proteins are imported via the translocase of the outer membrane (TOM) complex (Formosa et al. 2017). Some are laterally sorted into the inner mitochondrial membrane by the translocase of the inner membrane (TIM23) complex. Subunits bound for the matrix are then transported by the TIM23 complex, in association with the presequence translocase associated motor (PAM) (Sánchez-Caballero et al. 2016). The import across the mitochondrial inner membrane requires the proton gradient generated by the respiratory complexes. Other subunits, that sit on the IMS side of complex I, do not engage with TIM23, but rather undergo oxidative folding (Formosa et al. 2017).

Complex I is then assembled in a modular way, with different modules (e.g. N, Q, P) being assembled independently and then brought together in the mature complex (Figure 1.3). The modules themselves are made from discrete assembly intermediates. Assisting in this process are a number of helper proteins, called assembly factors, which are needed for assembly, but are not present on the mature complex (see next paragraph). The sequence of complex I assembly has been painstakingly unravelled over the last decade, since assembly intermediates were first identified in human patients (Triepels et al. 2001; Antonicka et al. 2003). The preferred method to study this process has historically been 1D-BN-PAGE and 2D-BN/SDS-PAGE, often followed by immunolabelling with specific antibodies. However, these approaches had three main drawbacks. First, studies were limited by available antibodies, meaning only a subset of proteins could be tracked. Second, migration profiles differed from lab to lab based on differences in solubilisation techniques used, often leading to confusingly contradictory molecular sizes given for assembly intermediates. Finally, by studying the assembly intermediates seen only in mutants, as mutants are by definition abnormal, the sequence of events may deviate from that seen in normal function. Recent advances in proteomics has allowed the advent of complexome profiling, which has revolutionised our ability to track the assembly of large numbers of proteins (Wessels et al. 2009; Heide et al. 2012; Wessels et al. 2013). The most extensive and up-to-date complexome profile on human complex I assembly was performed in 2016 (Guerrero-Castillo et al. 2016). Here the authors were able to track *de novo* assembly, in real time, in a wild-type setting.



**Figure 1.3: Complex I assembly in *Homo sapiens***

Schematic outlining the assembly of complex I. Complex I modules are made separately, and brought together in a particular order. Based on the model from Guerro-Castillo et al. 2016.

**Table 1.4: Complex I assembly factors in *Homo sapiens***

Gene name	Other gene names	Putative role in complex I assembly	Reference
ACAD9		Insertion of ND2	Nouws et al. 2010
ATP5SL			Stroud et al. 2016
DMAC1	TMEM261, C9orf123		Stroud et al. 2016
ECSIT		Insertion of ND2	Vogel et al. 2007b
FOXRED1	FP634		Calvo et al. 2010; Fassone et al. 2010; Formosa et al. 2015
NDUFAF1	CIA30, CGI-65	Insertion of ND2	Janssen et al. 2002; Vogel et al. 2005
NDUFAF2	NDUFA12L	Binding of N module	Ogilvie et al. 2005
NDUFAF3	C3orf60	Binding of Q with P <sub>p-a</sub>	Saada et al. 2009
NDUFAF4	C6orf66, HRPAP20	Binding of Q with P <sub>p-a</sub>	Saada et al. 2008
NDUFAF5	C20orf7	Hydroxylase activity	Sugiana et al. 2008
NDUFAF6	C8orf38	Squalene/ phytoene synthetase activity	Pagliarini et al. 2008
NDUFAF7	C2orf56, PRO1853	Methyltransferase activity	Carilla-Latorre et al. 2010
NDUFAF8	C17orf89		Floyd et al. 2016
NUBPL	C14orf127	FeS cluster insertion	Sheftel et al. 2009
TIMMDC1	C3orf1, UNQ247/PRO284	Insertion of ND1	Andrews et al. 2013
TMEM126B	HT007		Heide et al. 2012

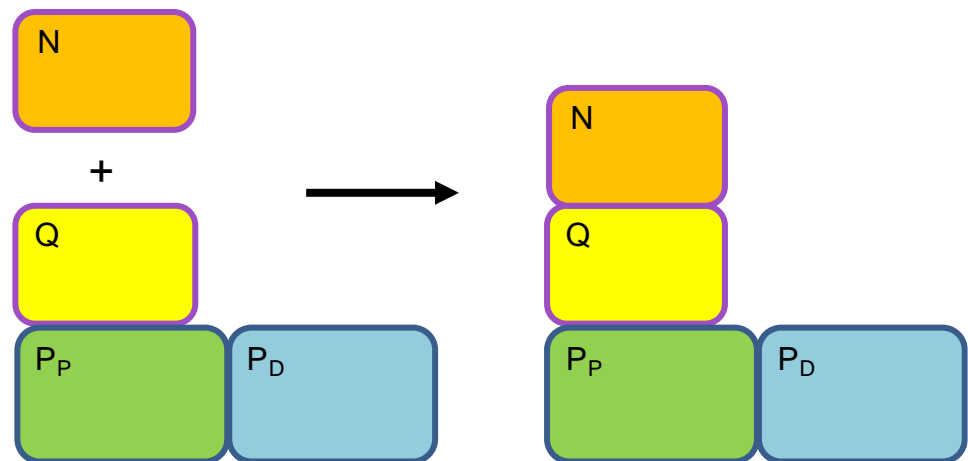
Treatment with chloramphenicol, which stopped synthesis of mtDNA encoded subunits, depleted existing respiratory complexes. This block was then removed, and the assembly of complex I tracked at time-points over a 24-hour period. This provided a model of unrivalled detail, which is the basis for Figure 1.3.

As mentioned, complex I assembly is helped by assembly factors. The first of these were identified in the yeast *Neurospora crassa*, in 1998, with the discovery of CIA30 and CIA84 (Küffner et al. 1998). Since then numerous other assembly factors have been identified. The number varies from group to group. In humans, there are currently 16 known assembly factors (Table 1.4) but this number increases every year as more are discovered (Stroud et al. 2016; Formosa et al. 2017). Many of these form tight interactions with complex I subunits, thus aiding in their detection and discovery. For example, there are nine assembly factors that bind to specific assembly intermediates (Andrews et al. 2013; Rhein et al. 2016). However, there are likely many others, which form only weak or transient interactions. It is probable that as analytical techniques grow more sensitive, more of these weak interacting assembly factors will be discovered. The roles of assembly factors vary. For example, some play structural, or scaffold, roles where they are needed to hold together or correctly insert subunits into the correct position, for example the assembly factor NDUFAF2 binds to an assembly intermediate of complex I, acting as a “place-holder”, until it is replaced by subunit NDUFA12 (Formosa et al. 2017). Others perform modifications to subunits, for example NDUFAF5 which adds a hydroxyl group to Arg-73 of the NDUF5 subunit (Rhein et al. 2016). An additional group of proteins exist which are implicated in complex I assembly but either have not been confirmed or are not exclusive complex I assembly factors. For example, TMEM70 was recently implicated in complex I assembly (Guerrero-Castillo et al. 2016), but is also a complex V assembly factor (Formosa et al. 2017).

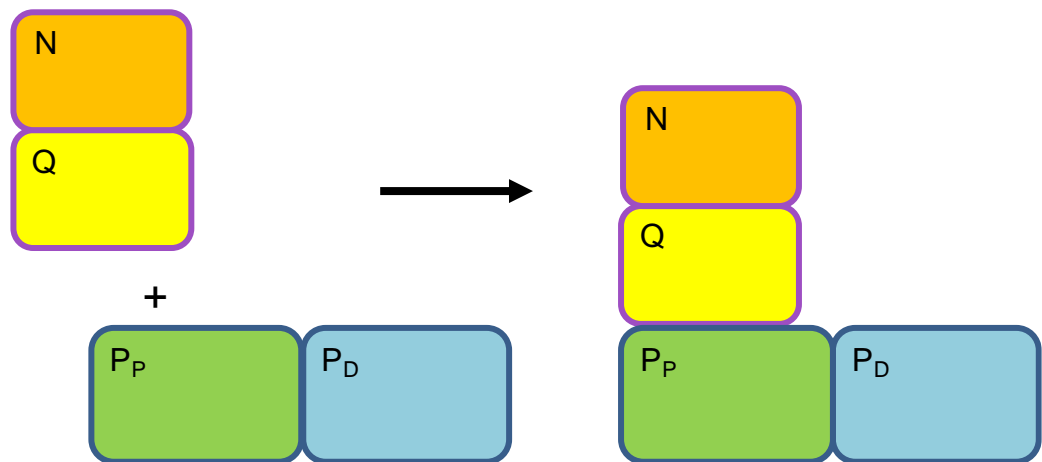
The process of complex I assembly is well characterised in humans, but varies in other groups. In fact, the sequence of events in mammals may be an evolutionary abnormality, as most other eukaryotes bring together the modules in a different order. For example, in algae, plants and yeast the N and Q module are brought together prior to attachment to the membrane arm (Vogel et al. 2007; Subrahmanian et al. 2016; Etienne Meyer, personal communication) (Figure 1.4). This means that an assembly defect seen in one group, may not translate well for another group, and so care should be taken in making comparisons.

Given that complex I assembly in the aerobic yeast *Yarrowia lipolytica* will be investigated in chapter 3, a brief outline of the assembly pathway in this organism will be

**A** *Homo sapiens*



**B** *Yarrowia lipolytica*



**Figure 1.4: Comparison of complex I assembly in *Homo sapiens* and *Yarrowia lipolytica***

Schematic outlining the sequence of events in assembly of complex I in **(A)** *Homo sapiens* and **(B)** *Yarrowia lipolytica*.

given (Figure 1.4). With no specific studies having been carried out in *Yarrowia lipolytica*, this pathway has to be inferred from fragmentary experimental observations of assembly intermediates formed in individual mutants and the process in other yeast species. Indeed, it is probably close to that seen in *Neurospora crassa*, which is quite different from that seen in humans (Videira & Duarte 2002; Vogel et al. 2007). The main difference is that, whilst in human the N-module is attached last, in yeast the N and Q module are first attached to each other, before being attached to the, possibly partial, membrane arm (Mimaki et al. 2012). However, care must be taken as a recent study on assembly in *Neurospora crassa* suggests that the process may be similar to mammals (Pereira et al. 2013). This highlights how little we currently know about complex I assembly in yeasts and non-mammalian model species. In addition, only two assembly factors have been studied in *Yarrowia lipolytica*, Ind1 and N7BML (Bych et al. 2008; Kmita et al. 2015), compared to 16 in humans.

## 1.4 Disease

### 1.4.1 Mitochondrial disorders

Mitochondrial disease is an umbrella term for a clinically heterogeneous set of disorders arising from mitochondrial dysfunction. Symptoms can affect single or multiple organs and can arise either in infancy or adulthood (Alston et al. 2017). Symptoms commonly include lactic acidosis, cardiomyopathy, leukodystrophy, skeletal myopathy, blindness and deafness (Fassone & Rahman 2012; Nouws et al. 2012; Vafai & Mootha 2012; Alston et al. 2017). Mitochondrial disorders are one of the most common inborn errors of metabolism (Calvo & Mootha 2010). Respiratory chain defects are estimated to affect approximately 1 in 5000 live births (Skladal et al. 2003). It is often said that these disorders predominantly affect tissues or organs with a high energy demand, such as CNS, heart, muscle, where mitochondria are more abundant (Alston et al. 2017; Craven et al. 2017). However, some caution that this is simplistic, and a more nuanced view takes into account nonlinear modes of pathogenesis and threshold effects (Vafai & Mootha 2012). Mitochondrial disorders often occur from respiratory chain dysfunction (Calvo & Mootha 2010) (see next paragraph), but can result in dysfunction of any mitochondrial pathway. As of the end of 2016, 281 genes (Craven et al. 2017), out of a total of 1,158 mitochondrial proteins (Mitocarta 2.0; Calvo et al. 2016), have mutations linked to human disease. Mitochondrial disorders can arise from mutations in either mtDNA or nuclear genes. The human mitochondrial genome is 16,569 bp and contains 37 genes, 13 of which encode proteins, structural subunits of the respiratory chain, as well as 2 ribosomal RNA and 22 tRNA genes. It was first sequenced in 1981 (Anderson



et al. 1981), and the first disease causing mutations were identified in 1988 (Holt et al. 1988; Wallace et al. 1988). Since then, pathogenic mutations have been found in every one of these genes (Alston et al. 2017). More common, given that the large majority of mitochondrial proteins are nuclear encoded, are nuclear gene mutations. The first of these were identified in 1995, where a patient with Leigh syndrome was found to have a mutation in *SDHA*, which encodes a subunit of complex II (Bourgeron et al. 1995). Since then over 250 have been discovered, with more being identified every year. Indeed, many genes with previously unknown functions in mitochondrial biology are being discovered in connection with disease, thereby adding to the inventory of disease causing genes.

### 1.4.2 Respiratory chain defects

Given the central role of mitochondria in cellular energy metabolism, it is unsurprising that the majority of mitochondrial disorders are caused by defects in the respiratory chain. Respiratory chain defects can result from mutations in structural subunits, but also in assembly factors, and in the numerous mitochondrial processes that affect respiratory chain function, for example mitochondrial DNA metabolism, protein import and cofactor biogenesis. As of 2016, pathogenic mutations have been found in 46 of the 90 respiratory chain structural subunit genes (Alston et al. 2017), as well as numerous others in assembly factors.

Complex I deficiency is the most common respiratory chain defect (Fassone & Rahman 2012; Mimaki et al. 2012; Rodenburg 2016) and is estimated to make up between a quarter and a third of all cases (Loeffen et al. 2000; Bugiani et al. 2004; Fassone & Rahman 2012). This may be because, as the largest respiratory complex, complex I has the largest probability of accruing pathogenic mutations (Rodenburg 2016). The first nuclear DNA mutation leading to complex I deficiency was in the *NDUFS8* subunits (Loeffen et al. 1998). Since then mutations have been found in genes coding for all the core subunits, and an increasing number in accessory subunits and assembly factors are found each year. As of 2017, disease causing mutations have been found in 26 of 44 structural subunits, and 10 of 16 assembly factors (Alston et al. 2017; Craven et al. 2017; Formosa et al. 2017; Dr. Charlotte Alston, personal communication). Clinical symptoms are similar to other mitochondrial disorders, making genotype-phenotype correlations hard, although the majority of complex I cases do have brain MRI abnormalities and 80% of paediatric cases develop Leigh syndrome (Koene et al. 2012; Fassone & Rahman 2012; Rodenburg 2016). However, for the most part, there are no symptoms specific for mutations in a single gene. The one exception, is in the complex

I assembly factor NUBPL, which will be the subject of chapter 3, where patients display distinct MRI patterns (Kevelam et al. 2013). Complex I deficiency is fatal in the majority of cases with 75% of cases fatal before 10 years old, and 50% before 2 years (Koene et al. 2012). Complex I deficiency has also been implicated in common disorders such as diabetes, Alzheimer's and Parkinson's disease (Nouws et al. 2012). Complex I deficiency was first implicated in the latter after it was noticed a high number of people using the illicit drug MPTP (1-methyl-4-phenyl-1,2,3,6- tetrahydropyridine), a known inhibitor of complex I, developed Parkinson's disease (Langston et al. 1983).

One of the recent research trends has been the mutually beneficial relationship between clinical and fundamental research in mitochondrial biology. Next generation sequencing technology has revolutionised the field (Craven et al. 2017) allowing new options which include whole genome sequencing, whole exome sequencing and sequencing a target panel of candidate genes (Alston et al. 2017). Indeed, whole genome or exome sequencing is now an achievable and affordable option for individual patients.

Discovery of new complex I disorders represents a good example of the new role technology plays, and highlights the innovative ways scientists have tackled the problem of identifying new disease causing mutations. For example, (Stroud et al. 2016) identified the assembly factor DMAC1, through CRISPR/Cas9 gene editing and profiling (Heide et al. 2012) using complexome profiling to identify the assembly factor TMEM126, and (Floyd et al. 2016) used proteomic profiling of protein-protein interactions to identify the assembly factor NDUFAF8.

Despite great advances in our understanding of mitochondrial diseases, treatment options are limited. Mitochondrial disorders are incurable and so must instead be managed. Future priorities must focus on better treatment options, or indeed a cure. Genome editing technologies may play a large part in this.

## 1.5 Evolution

### 1.5.1 Mitochondrial evolution and mitochondrion-related organelles

Mitochondrial evolution began with the endosymbiotic integration of an  $\alpha$ -proteobacteria into a eukaryotic cell (Margulis 1970; Margulis 1981; Gray et al. 1999). The exact timing, either early or late in eukaryotic evolution, and state of the engulfing eukaryotic cell has long been debated and remains a contentious issue (Poole et al. 2014; Pittis & Gabaldón 2016). However, currently opinion has centred around the fact that all extant eukaryotes have, or had, mitochondrially related organelles. The majority of eukaryotes possess

“canonical” mitochondria that engage in aerobic respiration, and so use OXPHOS to generate ATP, with oxygen as the final electron acceptor.

However there also exist non-canonical mitochondria. These are often the products of reductive evolution. Anaerobic mitochondria generate ATP using an electron transport chain and proton gradient, but do not need oxygen. Other terminal electron acceptors are used, for example fumarate (Tielens et al. 2002). There are also mitochondrion-related organelles (MROs) which evolved from mitochondria (Burki 2016). These are the product of reductive evolution and lack many features often associated with mitochondria, for example a mitochondrial genome, a membrane bound electron transport chain and invaginated inner mitochondrial membranes forming cristae. Their seemingly primitive appearance delayed appreciation that they were related to mitochondria. This led some to hypothesise that there existed a group of eukaryotes that lack mitochondria and diverged before mitochondria were acquired, the so-called Archezoa. The discovery of MROs meant that the Archezoa hypothesis is now considered false (Hjort et al. 2010). These MROs come in two main types. First, hydrogenosomes are simply defined as double-membrane bound organelles that produce hydrogen (Hjort et al. 2010). They oxidise pyruvate to acetate,  $H_2$  and  $CO_2$ , thus making ATP by substrate-level phosphorylation (Embley & Martin 2006). Hydrogenosomes have been found in many diverse groups including the Trichomonads, Chytridomycetes and Ciliates. Second, mitosomes have undergone even more reductive evolution than hydrogenosomes (Embley & Martin 2006). They are found in different groups, for example are present in *Giardia lamblia* and *Entamoeba histolytica*. Mitosomes were first identified in *Entamoeba histolytica* (Clark & Rogert 1995; Tovar et al. 1999). Mitosomes cannot synthesis ATP, but rather must import it. In addition, mitosomes have lost their genomes, as their individual genes became unnecessary, or were transferred to the nucleus.

Despite their structural and metabolic diversity, all mitochondria and MROs contained at least partial machinery for FeS cluster biosynthesis, the ISC pathway (see section 1.7). It was proposed that FeS biogenesis was the defining feature of this organelle (Lill 2009), and that no organism could lose the mitochondria as they are required for FeS clusters. It was therefore thought that all eukaryotes contain mitochondria or MROs. However, an exciting and unexpected discovery in 2016 changes this. An anaerobic microbe, *Monocercomonoides* sp., was found which apparently contained no mitochondria or MROs (Karnkowska et al. 2016). This raises the question as to how *Monocercomonoides* sp. produce FeS clusters, which are essential to life? It appears that they contain the alternative sulfur mobilisation (SUF) pathway, which here operates in the cytosol,

acquired through lateral gene transfer from bacteria. The SUF and the ISC pathway presumably co-existed in the *Monocercomonoides* sp. ancestor, creating redundancy and allowing the ISC pathway, along with the mitochondria, to be lost as they were no longer required. The intriguing example of *Monocercomonoides* sp. may be the exception that proves the rule. It reinforces the idea of FeS biosynthesis being the minimal function of the mitochondria, as only when this barrier is removed through redundancy, can mitochondria be lost (Burki 2016). It shows mitochondria themselves are not essential, but rather the biosynthetic pathways they contain.

### 1.5.2 Evolution of complex I

The evolution of respiratory complex I in eukaryotes is mainly the story of the evolution of the accessory subunits. Normally, bacterial complex I is made up of 14 subunits, the minimal enzyme, the same core 14 subunits found in all eukaryotes. For example, complex I in *E. coli* is made up of 13 subunits, although one of these is a fusion between two core subunits (NuoC and NuoD) (Berrisford et al. 2016). These subunits are necessary and sufficient for the enzymatic activity of complex I (Letts & Sazanov 2015). However, recently it has been shown that some bacterial complexes also contain accessory subunits. For example, complex I from *Thermus thermophilus* contains 16 subunits (Baradaran et al. 2013). Also, *Paracoccus denitrificans* complex I comprises 17 subunits, three of which are homologous to mammalian accessory subunits (B17.2, AQDQ/18, 13 kDa) (Yip et al. 2011). Given that *Paracoccus denitrificans* is an  $\alpha$ -proteobacteria, a close relative of the mitochondria's ancestor, it implies that the evolution of accessory subunits started prior to endosymbiosis.

However, despite the evolution of some accessory subunits in bacteria, accessory subunits are largely a eukaryotic innovation. To the minimal enzyme were added the accessory subunits, 30 in the case of humans, resulting in an approximate doubling in size of complex I (Wirth et al. 2016). 21 of the accessory subunits were evolved early in eukaryotic evolution, as they are present in all eukaryotic lineages (Gabaldón et al. 2005). The remaining subunits are lineage specific and evolved later. For example, there are 5 subunits in *Yarrowia lipolytica*, not present in the mammalian complex, and 8 subunits present in the mammalian complex not present in *Yarrowia lipolytica*. Another striking example of lineage specific complex I evolution is seen in plants, where complex I contains an extra domain: the carbonic anhydrase domain. This is attached to the membrane arm and extends into the mitochondrial matrix (Subrahmanian et al. 2016). It was first seen in single particle electron microscopy imaging of complex I in *Arabidopsis thaliana* (Dudkina et al. 2005). This domain contains five gamma-type carbonic

anhydrase proteins (Fromm et al. 2016b). Carbonic anhydrases are enzymes that catalyse a hydration reaction of  $\text{CO}_2$  to  $\text{HCO}_3^-$ , although no carbonic anhydrase activity has been shown yet for the complex I CA domain (Subrahmanian et al. 2016). Much of the variety seen in eukaryotic complex I is due to different complements of accessory subunits.

### 1.5.3 Loss of complex I

Despite lineage specific innovations, complex I structure, function and composition remains remarkably conserved across the eukaryotes. However, there are examples where complex I has been lost altogether. Some of these cases have occurred in anaerobic and parasitic eukaryotes, which have highly reduced mitochondria (see section 1.5.1). Five lineages of aerobic eukaryotes appear to have lost complex I in the course of evolution. These are the Cryptomycotan *Rozella allomyces* (James et al. 2013); the lineage which contains dinoflagellates, apicomplexans, *Chromera* and *Oxyrrhis* (Flegonotov et al. 2015), two groups of yeast (Gabaldón et al. 2005) and the parasitic plant genera, *Viscum* (Skippington et al. 2015; Petersen et al. 2015; Skippington et al. 2017). The first four examples are unicellular. However, *Viscum* is a multicellular plant species, where loss of complex I is unprecedented and surprising. It will be the focus of chapter 4, and more details can be found in the chapter introduction.

## 1.6 Cofactors

Mitochondria are commonly referred to as the “powerhouses of the cell”. This is a useful moniker, and still retains relevance. However, our current understanding of mitochondrial biology allows us to move beyond this name and look at mitochondrial function in its entirety (Pagliarini & Rutter 2013). The human mitochondrial proteome is currently thought to consist of 1,158 proteins (Mitocarta 2.0; Calvo et al. 2016). Of these only ~150 are directly involved in oxidative phosphorylation and ATP production. What do the remaining ~1000 proteins do? Many are involved with mtDNA metabolism, transport, homeostasis, and cofactor biosynthesis. Many of these will impact, indirectly, on OXPHOS, but many are quite separate. The last of these, cofactor biosynthesis, is one of the most important.

Cofactors, also referred to as prosthetic groups, are needed in biochemistry in order to augment the function of proteins. Amino acids are often chemically limited, and so cofactors give a protein biochemical versatility. Organic cofactors include nucleotides (for example FMN and FAD, already encountered in the respiratory chain), vitamins (for example biotin) and metal-organic compounds (for example molybdenum cofactor).

Inorganic cofactors include metal ions (for example  $\text{Zn}^{2+}$ ,  $\text{Cu}^{1+/2+}$ ,  $\text{Fe}^{2+/3+}$ ) and FeS clusters. Cofactors are especially important in redox reactions.

The respiratory chain contains many different cofactors (FeS, FMN, FAD, heme, copper). FeS clusters are present in complex I (8 clusters in humans), complex II (3 clusters in humans) and complex III (Rieske FeS subunit). Therefore, in understanding complex I, II and III assembly and regulation it is important to understand the biosynthesis of FeS clusters, which will be the focus of the remainder of this introduction.

## 1.7 Iron-sulfur clusters

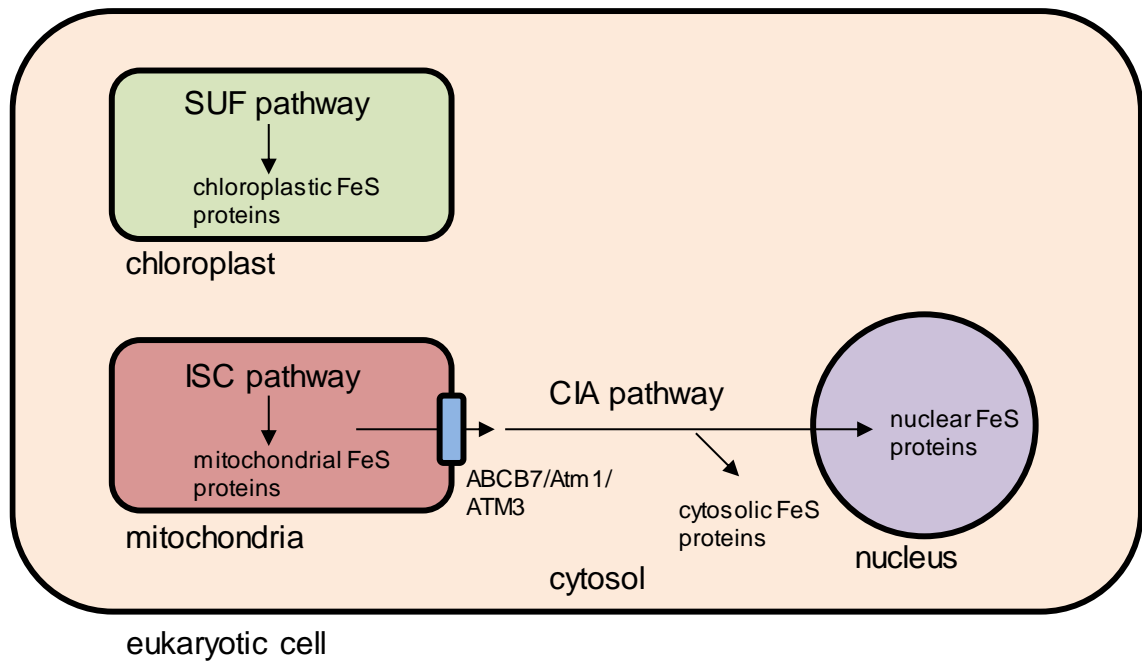
### 1.7.1 Iron-sulfur cluster function

FeS clusters were first discovered in the 1960s (Rouault 2015). They are excellent acceptors and donors of electrons, with redox potentials between -500 mV and +300 mV, and so are often found in redox enzymes (Lill 2009). Indeed, complex I is the largest FeS cluster containing enzyme known (Beinert 2000). They also have roles as catalysts in enzymatic reactions, in sensing intracellular or environmental conditions to regulate gene expression, and in stabilisation of protein structure (Lill & Muhlenhoff 2008). They are thought to be one of the most ancient cofactors, and may play a role in the early evolution of life (Russell & Martin 2004). FeS clusters come in many forms but the most common are rhombic  $[\text{2Fe-2S}]^{1+/2+}$  and cubane  $[\text{4Fe-4S}]^{1+/2+}$ . Interconversion of cluster types has been observed, demonstrating the modular nature of the clusters (Beinert et al. 1997). Cubane clusters are less stable than rhombic ones, and FeS clusters are generally easily destabilised in the presence of oxygen. FeS clusters are usually coordinated by cysteine residues, but can also be bound by histidine, arginine, serine, and other small molecules (Lill & Muhlenhoff 2008). For example, in the Rieske FeS cluster in complex III, one of the Fe atoms is coordinated by two histidine residues (Maio & Rouault 2015). The number of FeS proteins in a cell is hard to estimate, but they participate in many essential cellular processes, including DNA metabolism and photosynthesis in plants.

### 1.7.2 Iron-sulfur cluster assembly

After their initial discovery in the 1960's, *in vitro* work showed that FeS clusters could assemble spontaneously on proteins, through chemical reconstitution (Lill 2009). However, work in the 1990s showed that the process *in vivo* required catalysis and required the involvement of a large number of proteins (Lill & Muhlenhoff 2008). Our

understanding of the assembly of FeS clusters has advanced greatly in the last two decades, with mitochondria playing a key role in this.



**Figure 1.5: FeS assembly pathways in the eukaryotic cell**

Schematic showing the FeS assembly pathways found in eukaryotic cells. The SUF pathway is predominantly found in photosynthetic eukaryotes. The connection between the mitochondrial ISC and cytosolic CIA pathway is mediated by the ABC transporter ABCB7/Atm1/ATM3.

FeS cluster assembly has been extensively studied in bacteria, especially *Escherichia coli* and the nitrogen fixing (azototrophic) *Azotobacter vinelandii*. Three distinct pathways for FeS biogenesis have been identified (Lill 2009; Rouault & Maio 2017): the NIF (nitrogen fixation), which is specific for the maturation of FeS clusters in the enzyme nitrogenase; the ISC (iron-sulfur cluster) and the SUF (sulfur-mobilisation) pathway. The ISC and the SUF pathway were transferred during endosymbiosis and can now be found in eukaryotes, along with a third pathway, the CIA (cytosolic FeS protein assembly) pathway (Lill & Muhlenhoff 2008). The ISC pathway is found in the mitochondria, and provides FeS clusters for mitochondria proteins, including the respiratory chain (Figure 1.5). The SUF pathway is found in plastids, for example the chloroplasts in plants, and so is only present in some eukaryotic lineages. The SUF pathway provides the FeS clusters needed for plastid processes, for example photosynthesis. The eukaryotic innovation, the CIA pathway, is needed for FeS clusters used in the cytosol and the nucleus. The CIA pathway requires the mitochondrial ISC pathway, and is connected to it by the mitochondrial ISC export system, made up of the ABC transporter ABCB7/Atm1/ATM3.

The generation of an FeS cluster in any of these pathways, can be broken down into two basic steps (Lill 2009; Maio & Rouault 2015): 1) *de novo* assembly of the nascent FeS cluster and 2) transfer of the FeS cluster to recipient apoproteins. During step 1 a cysteine desulphurase acquires a protein-bound persulfide ( $S^0$ ) from cysteine which is then reduced to a sulfide ( $S^{2-}$ ) before being transferred to a scaffold protein and is combined with Fe, to form the nascent FeS cluster. The scaffold protein provides a platform for this reaction to take place and contains cysteine residues to bind the FeS cluster in a labile manner. In step 2 this labile FeS cluster is then transferred to recipient enzymes with the help of numerous carrier proteins, thereby converting apoproteins to holoproteins. These FeS carriers, also called transfer proteins or intermediate scaffolds, ensure correct dissociation of the nascent FeS cluster from the scaffold, correct targeting to a specific recipient enzyme, and then correct transfer of the FeS cluster onto the acceptor site of the recipient enzyme. The details of the second step are less well understood than the first step. These steps are common to all pathways, although the protein machineries involved can be quite different. Here I will focus on the mitochondrial ISC pathway.

### 1.7.3 The mitochondrial ISC pathway

The mitochondrial ISC pathway is required for mitochondrial, and also cytosolic and nuclear, FeS proteins. The pathway has been steadily elucidated over the past decade,



predominantly in the model yeast, *Saccharomyces cerevisiae*, but also in human cell cultures. The details differ quite substantially between *Saccharomyces cerevisiae* and mammals, making a unified eukaryotic model hard to create. I will first briefly outline the model in *Saccharomyces cerevisiae*.

The process can broadly be split into four steps (Braymer & Lill 2017) (an expansion on the two basic steps outlined in section 1.72). In stage 1, [2Fe-2S] are synthesised *de novo* on the scaffold protein Isu1, in complex with the cysteine desulfurase, Nfs1, and two helper proteins Isd11 and Acp1. The ferredoxin Yah1 as well as Yhf1 (frataxin in humans), which has been proposed as an iron donor or regulator, are involved at this stage as well. In stage 2, Isu1 releases the [2Fe-2S] cluster, which is then taken by the monothiol glutaredoxin, Grx5 for downstream trafficking. The release of the cluster from Isu1 is helped by two chaperone proteins, Ssq1 (a Hsp70 protein) and Jac1. Grx5 then either directs the cluster to a target [2Fe-2S] protein, feeds it into late ISC machinery for [4Fe-4S] cluster synthesis, or, in this model, directs the cluster to form a sulfur intermediate (X-S) for export to the cytosol and the involvement in the CIA pathway. Additionally, the [2Fe-2S] cluster may be directed to target proteins before involvement with Grx5. In stage 3, [4Fe-4S] clusters are synthesised by the late ISC machinery, combining [2Fe-2S] clusters from Grx5. This is achieved by a complex of proteins consisting of Iba57, Isa1 and Isa2. In stage 4, [4Fe-4S] clusters are trafficked and inserted into target [4Fe-4S] proteins, facilitated by ISC targeting factors. These include Bol1, Bol3, Nfu1 and Ind1. Note, the role of the plant homologs of Grx5 (GRXS15), Isu1 (ISU1) and Nfu1 (NFU4 and NFU5) will be investigated in chapter 5. The human homolog of Ind1 is NUBPL, and has already been mentioned in the complex I assembly section. It will be the focus of chapter 3. Some inconsistencies exist in the model. For example, Bol1 and Bol3 have been shown to interact with Grx5, from an earlier part in the process (Melber et al. 2016; Uzarska et al. 2016), which would suggest that either Grx5 plays an additional role in the late ISC machinery, or that the Bol proteins also affect [2Fe-2S] cluster synthesis.

Indeed, the model for the mammalian ISC pathway differs in many respects. Some of these are due to lineage specific differences, but in other ways it is an alternative model of FeS cluster assembly. The main difference is, in the mammalian model, the complex of HSC20-HSPA9-ISCU1 (Isu1-Ssq1-Jac1 in yeast) plays a more central role (Maio & Rouault 2015; Rouault & Maio 2017). This complex is responsible for inserting FeS clusters into almost all recipient proteins, with FeS carrier proteins playing a more marginal role. Here, the transfer of [2Fe-2S] and [4Fe-4S] clusters are not separate. Another difference is that there is evidence of an additional cytosolic localisation of many

ISC components. It is therefore proposed that the CIA pathway plays a less important role, and that the cytosolic ISC pathway supplies FeS clusters for nuclear and cytosolic FeS proteins (Rouault 2012; Maio & Rouault 2015). The differences between the two models will be more fully addressed in the discussion.

### **1.7.4 Iron-sulfur assembly in plants**

The FeS assembly pathways seen in plants are broadly similar to those seen in yeast and humans, except with the addition of the SUF pathway in the chloroplasts (Couturier et al. 2013; Balk & Schaedler 2014). Plants provide a good, but underexplored, study system for understanding FeS assembly. It provides an opportunity to address inconsistencies in the current models. For example, Nfu1 is essential in humans but yeast mutants only have a mild phenotype. Studying Nfu in another eukaryotic model, like plants, is therefore valuable. Also, given that the main model, *Saccharomyces cerevisiae*, lacks complex I, it provides a good opportunity to explore FeS provision for respiratory complexes.

Recent work in plants has focused on the CIA pathway, for example looking at NBP35 (Bastow et al. 2017) and AE7 (Luo et al. 2012), as well as the connection between the cytosolic and mitochondrial pathways (Bernard et al. 2009; Schaedler et al. 2014). The pathways are connected by the ABC half-transporter, ATM3, and studies have helped identify the previously unknown exported sulfur compound (X-S) as glutathione polysulfide.

Apart from a recent study on the *Arabidopsis thaliana* Ind1 homolog, INDH (Wydro et al. 2013), and two studies of the Grx5 homolog, GRXS15 (Moseler et al. 2015; Stroehler et al. 2015), relatively little work has been done on the mitochondrial ISC pathway in plants in recent years. Homologs of the main ISC components are found but the pathway is less well resolved. One difference, for example, is that there are two NFU proteins, NFU4 and NFU5, as opposed to the one found in yeast and humans. Also, there are three versions of Isu1 (ISU1, ISU2 and ISU3). These differences, and the poor resolution of the pathway provide a good opportunity to explore the functions of late acting FeS carrier proteins in the *Arabidopsis thaliana* model system.

### **1.7.5 Iron-sulfur assembly and human disease**

Mitochondrial disorders, which have been covered in section 1.41, are often caused by dysfunction in the respiratory chain. Recent findings have shown that defects in cofactor biosynthesis also contribute to disease. This can affect indirectly the respiratory chain, but also enzymes used by other mitochondrial metabolic systems.

Numerous human patients have been identified with defects in FeS assembly, especially in the mitochondrial ISC pathway. As of the end of 2016, 11 genes in the ISC pathway have been discovered with pathogenic mutations (*BOLA3*, *FDX1L*, *FXN*, *GLRX5*, *IBA57*, *ISCA2*, *ISCU*, *LYRM4*, *NFU1*, *NFS1* and *NUBPL*) (Craven et al. 2017). Of these, *NUBPL* will be investigated in chapter 3, while the plant homologs of *ISCU*, *NFU1* and *GLRX5* will be investigated in chapter 5.

As well as OXPHOS disorders, specific syndromes have been associated with mitochondrial ISC defects, for example sideroblastic anemia associated with a mutation in *GLRX5* and Friedreichs ataxia, associated with mutations in *FXN* (Stehling et al. 2014). Given the connection between the ISC and CIA pathway, defects in the ISC pathway can affect nuclear FeS proteins involved with DNA metabolism and so can also be linked to cancer (Sheftel et al. 2012). Studying the diseases resulting in ISC dysfunction can prove valuable in understanding their function.

### 1.8 Objectives

Mitochondrial research is once again a subject of central interest. Great advances have been made in the fields of complex I structure, function and assembly, mitochondrial disorders, mitochondrial evolution and FeS cluster biogenesis. This thesis aims to contribute to all these areas.

First, in chapter 3, the role of the complex I assembly factor *NUBPL* in complex I assembly and disease will be investigated. This will have relevance for clinicians, but will also contribute to our understanding of its role in assembly. Patient mutations in *NUBPL* will be recreated and studied in the model yeast *Yarrowia lipolytica*.

Second, in chapter 4, the evolutionary mystery of the loss of complex I in the multicellular plant species *Viscum album* will be investigated using proteomic and biochemical techniques. If it has indeed lost complex I, it will raise questions about whether complex I is essential for multicellular aerobic life.

Thirdly, in chapter 5, the FeS assembly pathway in plant mitochondria will be investigated to better understand cofactor provision for the mitochondria and respiratory complexes. The roles of the FeS carrier proteins, *NFU4* and *NFU5*, will be studied for the first time in *Arabidopsis thaliana*.

These three topics will enhance our understanding of the function and evolution of complex I and mitochondria. By studying mechanistic details, but then setting them in the context of evolution a fuller understanding can be achieved.

## 2 Materials and methods

### 2.1 Bioinformatics

#### 2.1.1 Databases

<i>Arabidopsis thaliana</i>	<a href="http://www.arabidopsis.org">http://www.arabidopsis.org</a>
<i>Yarrowia lipolytica</i>	<a href="http://www.genome.jp/kegg-bin/show_organism?org=yli">http://www.genome.jp/kegg-bin/show_organism?org=yli</a> <a href="http://fungi.ensembl.org/Yarrowia_lipolytica/Info/Index">http://fungi.ensembl.org/Yarrowia_lipolytica/Info/Index</a>
ExAC	<a href="http://exac.broadinstitute.org/">http://exac.broadinstitute.org/</a>
Protein sequences	<a href="http://www.uniprot.org/">http://www.uniprot.org/</a>

#### 2.1.2 Alignments

Nucleic and amino acid alignments were performed using ClustalOmega (Sievers et al. 2011), hosted at <http://www.ebi.ac.uk/Tools/msa/clustalo/>. Alignments were coloured using BoxShade v.3.21 (Kay Hoffman and Michael Baron), hosted at [http://www.ch.embnet.org/software/BOX\\_form.html](http://www.ch.embnet.org/software/BOX_form.html).

#### 2.1.3 Homology modelling

Homology modelling was performed using the INTFold server (McGuffin et al. 2015), hosted at <http://www.reading.ac.uk/bioinf/IntFOLD/>. Structures were visualised using CCP4mg (McNicholas et al. 2011).

#### 2.1.4 Illustration of complex I structure

Complete structure of complex I from *Bos taurus*, Protein Data Bank code: 5LDW, was visualised and illustrated with the UCSF Chimera package (Pettersen et al. 2004).

#### 2.1.5 Prediction of functional effects of human nsSNPs

Prediction of the functional effects of human nsSNPs was performed with the PolyPhen-2 tool (Adzhubei et al. 2010), hosted at <http://genetics.bwh.harvard.edu/pph2/>.

### 2.2 Chemicals

Chemicals were obtained from: Alfa Aesar (Heysham, UK), Amersham (LittleChalfont, UK), AMSBio (Abingdon, UK), Anachem (Luton, UK), Bio-Rad Laboratories Ltd (Hemel, Hempstead, UK), Biochem (Cosne-Cours-sur-Loire, France), Cayman Chemical (USA), DBC Foodservice (Petersfield, UK), Duchefa (Haarlem, Netherlands), Expedeon

(Swavesey, UK), Formedium (Hunstanton, UK), Melford (Ipswich, UK), Qiagen (Manchester, UK), Roche (Burgess Hill, UK), Sigma-Aldrich/Fluka (Dorset, UK), Thermo Fisher Scientific (Loughborough, UK), VWR/BDH (Lutterworth, UK).

## 2.3 Organisms

### 2.3.1 *Arabidopsis thaliana*

**Table 2.1: List of *Arabidopsis thaliana* lines**

Line	Gene ID	Background	Stock centre ID	Reference
<i>grxs15-3</i> UBQ10:GRXS15 K83A	AT3G15660	Col-0	N/A	Moseler et al. 2015
<i>grxs15</i> amiRNA	AT3G15660	Col-0	N/A	Stroher et al. 2015
<i>indh</i>	AT4G19540	Col-0	GK_956A05	Wydro et al. 2013
<i>isu1-1</i>	AT4G22220	Col-0	SALK_006332	N/A
<i>ndufs4</i>	AT5G67590	Col-0	SAIL_596_E11	Meyer et al. 2009
<i>nfu4-1</i>	AT3G20970	Col-0	SALK_035493	N/A
<i>nfu4-2</i>	AT3G20970	Col-0	SALK_061018	N/A
<i>nfu4-4</i>	AT3G20970	Col-0	SAIL_1233_C08	N/A
<i>nfu5-1</i>	AT1G51390	Col-0	WiscDsLoxHs069_06B	N/A
<i>nfu5-2</i>	AT1G51390	Col-4	SK24394	N/A
<i>nfu5-3</i>	AT1G51390	Ler	GT_3_2834	N/A
<i>rug3-1</i>	AT5G60870	Col-0	SALK_092071	Kühn et al. 2011

### 2.3.2 *Yarrowia lipolytica*

**Table 2.2: List of *Yarrowia lipolytica* strains**

Strain	Background	Reference
<i>ind1Δ</i>	GB10	Bych et al. 2008
<i>ind1Δ</i> + IND1-N272QfsX31	GB10	Wydro et al. 2013
<i>nubmΔ</i>	GB10	Ulrich Brandt
<i>nucmΔ</i>	GB10	Ulrich Brandt
<i>nucmΔ</i> + NUCM-Y144F	GB10	Tocilescu et al. 2010
<i>nukmΔ</i>	GB10	Ulrich Brandt

GB10 strain: *ura3-302*, *leu2-270*, *lys11-23*, *NUGN-Htg2*, *NDH2i*, *MatB*.

### 2.3.3 *Escherichia coli*

**Table 2.3: List of *Escherichia coli* strains**

Strain	Antibiotic sensitivity	Genotype
DH5 $\alpha$	Carbenicillin	<i>F</i> – $\phi$ 80 <i>lacZ</i> $\Delta$ M15 $\Delta$ ( <i>lacZYA-argF</i> ) U169 <i>recA1 endA1 hsdR17</i> ( <i>rK</i> –, <i>mK</i> +) <i>phoA supE44</i> $\lambda$ – <i>thi-1 gyrA96 relA1</i>
Top10	Carbenicillin	<i>F</i> – <i>mcrA</i> $\Delta$ ( <i>mrr-hsdRMS-mcrBC</i> ) $\phi$ 80 <i>lacZ</i> $\Delta$ M15 $\Delta$ <i>lacX74 recA1 araD139</i> $\Delta$ ( <i>ara leu</i> ) 7697 <i>galU galK rpsL</i> ( <i>StrR</i> ) <i>endA1 nupG</i>
Rosetta +pLYS +pISC	Carbenicillin, chloramphenicol, tetracycline	<i>F ompT hsdSB</i> ( <i>rB</i> – <i>mB</i> –) <i>gal dcm</i> (DE3) <i>pRARE</i> ( <i>Cam</i> <sup>R</sup> )

## 2.4 Molecular techniques: Polymerase Chain Reaction

### 2.4.1 DNA extraction from *Arabidopsis thaliana*

DNA was extracted from *Arabidopsis thaliana* leaf tissue as described in (Kasajima et al. 2004). Briefly, to leaf tissue, 200  $\mu$ l of buffer (Edwards solution: 200 mM Tris-HCL, pH 7.5, 250 mM NaCl, 25 mM EDTA, 0.5% (w/v) SDS (Edwards et al. 1991), diluted 1:10 in TE buffer (10 mM Tris-HCL, pH 8.0, 1 mM EDTA) was added, and then the tissue ground with a mini-pestle in a 1.5 ml tube. Debris was removed by centrifugation at 16,100 x *g* for 5 minutes. 1  $\mu$ l of supernatant was then used for genotyping PCR.

### 2.4.2 DNA extraction from *Yarrowia lipolytica*

*Yarrowia lipolytica* cells were collected by centrifugation at 3,500 x *g* and resuspended in 0.9 ml solution 1 (0.9 M sorbitol, 0.1 M EDTA, 0.01 M KPi-buffer, pH 7.4). 0.1 ml of solution 2 was added (2 mg ml<sup>–1</sup> zymolyase, 10  $\mu$  ml<sup>–1</sup>  $\beta$ -mercaptoethanol), mixed, and incubated for 30 minutes at 37°C. Cells were collected by centrifugation and resuspended in 0.4 ml TE (10 mM Tris-HCL, pH 8.0, 1 mM EDTA). 30  $\mu$ l TES was added and incubated for 30 minutes at 60°C. 80  $\mu$ l 5 M potassium acetate was added and incubated on ice for 1 hour. After centrifugation at 16,100 x *g*, 1 ml ethanol was added to the supernatant. After another round of centrifugation, the DNA was washed in 70% (v/v) ethanol, and resuspended in TE buffer.

### **2.4.3 RNA extraction from *Arabidopsis thaliana***

RNA extraction from *Arabidopsis thaliana* was performed using a RNeasy® Plant Mini Kit (Qiagen) according to the manufacturer's instructions.

### **2.4.4 RNA extraction from *Yarrowia lipolytica***

RNA extraction from *Yarrowia lipolytica* was performed using a RNeasy® Mini Kit (Qiagen) according to the manufacturer's instructions.

### **2.4.5 cDNA synthesis and reverse transcription**

For *Arabidopsis thaliana* and *Yarrowia lipolytica* RT-PCR, extracted RNA was converted to cDNA by reverse transcription. Briefly, between 0.7 and 2.4 µg of RNA were used (equal amounts taken for each sample set). RNA was mixed with 0.7 µl dT (20)VN primer, and made up to 12 µl with dH<sub>2</sub>O. Samples were incubated for 5 minutes at 70°C, then cooled on ice. 4 µl of 5xRT buffer (M-MLV Reverse Transcriptase kit, Invitrogen), 2 µl 10 mM dNTPs and 1 µl RNase OUT (Invitrogen) were added to samples, which were then incubated for 5 minutes at 37°C. 1 µl reverse transcriptase (M-MVL Reverse Transcriptase kit, Invitrogen) was added, and then samples were incubated for 60 minutes at 42°C followed by 10 minutes at 70°C. 30 µl dH<sub>2</sub>O were added to give a final volume of 50 µl. 1 µl was then used as a template for polymerase chain reaction (see section 2.4.6).

### **2.4.6 Polymerase chain reaction**

Polymerase chain reactions were all performed on Eppendorf Nexus thermocyclers. Reactions were made up with polymerase, an appropriate polymerase buffer, dNTPs, oligonucleotides and a DNA template. Different polymerases were used for different types of PCR, for example for genotyping PCR, a lab home-brew stock of EcoTaq polymerase and for cloning Phusion® High-Fidelity DNA polymerase (NEB). Thermocycling protocols were adjusted for different polymerases. Denaturation steps were performed at 94 °C (Eco-Taq) or 98 °C (Phusion), annealing temperatures were adjusted to be appropriate for particular oligonucleotides, and elongation steps were performed at 72 °C. Typically between 30-35 thermocycles were performed. Touchdown PCR was performed for certain genotyping reactions.

### **2.4.7 DNA extraction from agarose gel**

DNA was extracted from agarose gel slices using a QIAquick® PCR purification kit (Qiagen) according to the manufacturer's instructions.

### 2.4.8 Agarose gel electrophoresis

DNA fragments were separated in a 1% (w/v) agarose gels, containing 0.04 % (w/v) ethidium bromide. Gels were electrophoresed at 100 V, in TAE buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA). DNA bands were visualised in a UV transilluminator (G:Box, Syngene). 100 bp and 1 kb ladders (NewEngland Biolabs) were used to estimate DNA band size.

### 2.4.9 Oligonucleotides

Oligonucleotides were obtained from Eurofins MWG Operon and were used at a concentration of either 10 or 25 µM.

**Table 2.4: Primers used for RT-PCR**

Oligonucleotide	Gene	5'-3' sequence	Species
IND1-RT-F	<i>IND1</i>	CGTCTCCGTGAACACAGCAC	<i>Yarrowia lipolytica</i>
IND1-RT-R		CACGCCACTGCCTTGTTAG	
ACT1-RT-F	<i>ACT1</i>	TATGTGCAAGGCCGGTTTCG	<i>Yarrowia lipolytica</i>
ACT1-RT-R		TCGATGGGGTATCGGAGGGT	
PYC1-RT-F	<i>PYC1</i>	TTCCTGAGACCAAGGTGGATC	<i>Yarrowia lipolytica</i>
PYC1-RT-R		CTTGAATCGGTGCATAGAGAGCC	
NFU4-RT-F	<i>NFU4</i>	AGGTCGCAAGGTCGTTTCTG	<i>Arabidopsis thaliana</i>
NFU4 -RT-R		AGCCCAGAAACAACACCAAG	
NFU5-RT-F	<i>NFU5</i>	AGCTTCCGTAAC TTGTTCTCC	<i>Arabidopsis thaliana</i>
NFU5-RT-R		AGATGTTCCCTCCTCTTCAC	
ACT2-RT-F	<i>ACT2</i>	CCCAAAGGCCAACAGAGAGA	<i>Arabidopsis thaliana</i>
ACT2-RT-R		ACCATCACCAGAATCCAGCA	



**Table 2.5: Primers used for genotyping**

Mutant	Genotype	Oligonucleotide	5'-3' sequence
<i>nfu4-1</i>	Wild type	<i>nfu4-1</i> LB	CCCAGTTGCTTTAATGAGCTG
		<i>nfu4-1</i> RB	AACTGTACGCCAAACGTGAAC
	T-DNA	LBb1	GCGTGGACCGCTTGCTGCAACT
		<i>nfu4-1</i> RB	
<i>nfu4-2</i>	Wild type	<i>nfu4-2</i> LB	TCTCCAATCCCTCAACAAATG
		<i>nfu4-2</i> RB	TCTGACGAAAACAATCCATCC
	T-DNA	LBb1	GCGTGGACCGCTTGCTGCAACT
		<i>nfu4-2</i> RB	
<i>nfu4-4</i>	Wild type	<i>nfu4-4</i> LB	TCTTGATTCTCCAATGGTTGC
		<i>nfu4-4</i> RB	CAAGCCTGAGATATTTGCAGC
	T-DNA	LB3	TAGCATCTGAATTTTCATAACCAATCTCGATACAC
		<i>nfu4-4</i> RB	
<i>nfu5-1</i>	Wild type	<i>nfu5-1</i> LB	TTCACCATCGAATTCTTGCTC
		<i>nfu5-1</i> RB	CGTCGCTCTCTGTTTATCTCG
	T-DNA	L4	TGATCCATGTAGATTTCCCGGACATGAAG
		<i>nfu5-1</i> RB	
<i>nfu5-2</i>	Wild type	<i>nfu5-2</i> LB	TTCTTAAGCTCTGGTCCCTCC
		<i>nfu5-2</i> RB	CGAGATAAACAGAGAGCGACG
	T-DNA	pSKTAIL-L3	ATACGACGGATCGTAATTTGTCTG
		<i>nfu5-2</i> RB	
<i>nfu5-3</i>	Wild type	<i>nfu5-3</i> LB	ACAGATGAACTTGGACAACCG
		<i>nfu5-3</i> RB	CGTCGCTCTCTGTTTATCTCG
	T-DNA	Ds3-1	ACCCGACCGGATCGTATCGGT
		<i>nfu5-3</i> RB	
<i>isu1-1</i>	Wild type	<i>isu1-1</i> F1	GAACCATCTAAACCGTCCACGT
		<i>isu1-1</i> R1	GGGTGAAAGGCAAAGCTATGGAG
	T-DNA	LBb1.3	ATTTTGCCGATTTTCGGAAC
		<i>isu1-1</i> R1	

**Table 2.6: Primers used for mutagenesis**

Amino acid change	DNA change	Oligonucleotide	5'-3' sequence
L102P	t305c	L102P-F	CTGAGGGTAGGACTCCCCGATGTGGACATCTTC
		L102P-R	GAAGATGTCCACATCGGGGAGTCCTACCCCTCAG
D103Y	g307t	D103Y-F	CCGAAGATGTCCACATAGAGGAGTCCTACCCTC
		D103Y-R	GAGGGTAGGACTCCTCTATGTGGACATCTTCGG
G136D	g407a, g408t	G136D-F	CCCATGCTCATGACCTGTATATCGAATTCGACATTGGAATC
		G136D-R	GATTCCAATGTGCGAAATTCGATATACAGGTCATGAGCATGGG
M180V	a538g	M180V-F	GGGTGGCAAGTCCACCACCAGAACGTCCA
		M180V-R	TGGACGTTCTGGTGGTGGACTTGCCACCC
M180A	a538g, t539c	M180A-F	CCGGGTGGCAAGTCCGCCACCAGAACGTCCAA
		M180A-R	TTGGACGTTCTGGTGGCGGACTTGCCACCCGG
L191F	c571t, g573c	L191F-F	AGTTTGGGCAATAGTGAATTGAACGTCTCCCGTTCCGG
		L191F-R	CCGGAACGGGAGACGTTCAATCACTATTGCCCAAAC
G285C	g853t	G285C-F	GCTCACAGTCAGACAAGTGTGTTCTGTTGCTGTG
		G285C-R	CACAGCAACAGGAACACACTTGTCTGACTGTGAGC

**Table 2.7: Primers used for cloning and sequencing**

Oligonucleotide	Purpose	5'-3' sequence
AM1	Cloning of <i>IND1</i>	TGCCCATGCTTCTACACCTCC
AM2	Cloning of <i>IND1</i>	TGCTGCCGTCGCTGGTGCCTTT
AM27	Cloning Ind1 protein expression vectors	GGCCATGGAAAACCCCTGGGTATC
AM28	Cloning Ind1 protein expression vectors	GGCTCGAGCTATTTTTCAAATTGAGG
KB86	Sequencing Ind1-strep-pUB26 and variants	CTGCAGGTACACCCAACTGCTCAGC
KB87	Sequencing Ind1-strep-pUB26 and variants	GGCATATGGAAAACCCCTGGGTATC

## **2.5 Molecular techniques: Plasmid construction**

### **2.5.1 Restriction digest**

Restriction digest was performed according to manufacturer's instructions (mainly New England Biolabs).

### **2.5.2 Plasmid isolation**

Plasmid isolation was performed using a QIAprep® Spin Miniprep kit (QIAGEN), according to the manufacturer's instructions.

### **2.5.3 DNA sequencing**

Sequencing was performed by Source Bioscience and Eurofins MWG Operon. For sequencing of pGEM®-T Easy vectors, primers SP6 and T7 were used and provided by sequencing company.

### **2.5.4 Ligation**

Ligation was performed using T4 ligase, according to manufacturer's instructions.

### **2.5.5 Mutagenesis**

For *Yarrowia lipolytica* vectors, *IND1*-strep sequence was isolated from the pUB4-*IND1*-strep (Bych et al. 2008, also, called pKB73 by Dr. Katrine Bych. In table 2.8 as pAM3) by PCR, using primers AM1 and AM2. In the last three cycles the mixture was spiked with Taq-polymerase, to A-tail the PCR product. The PCR product was inserted into an pGEM®-T Easy vector (Promega), according to the manufacturer's instructions, to make plasmid pAM1. Mutagenesis primers, table 2.6, were designed using the online QuikChange primer design tool, hosted at <http://www.genomics.agilent.com/primerDesignProgram.jsp>. pAM1 was used as a template for site-directed mutagenesis PCR, which was carried out using a Quikchange® II mutagenesis kit (Agilent), as per the manufacturer's instructions. Mutagenesis was confirmed by sequencing. The mutated *IND1*-strep sequence was then subcloned out of the pGEM®-T Easy vector, and inserted into pUB4-*IND1*-strep, between XbaI and NsiI restriction sites. Plasmids were confirmed by sequencing and diagnostic restriction digest.

### **2.5.6 Protein expression vectors**

For protein expression in *E. coli*, *IND1*-strep, excluding the first 36 amino acids, in the pET15b vector (Novagen) was used (Bych et al. 2008, also, called pKB71 by Dr. Katrine Bych). For expressing Ind1-L102P and Ind1-D103Y, sequence was isolated from the

*Yarrowia lipolytica* vectors, pAM8 and pAM13, by PCR, using primers AM27 and AM28. These sequences were then inserted into the pKB71 vector, between the NcoI and XhoI restriction sites. Plasmids were confirmed by sequencing and diagnostic restriction digest. Plasmids expressing Ind1-G80S and Ind1-K81Q had previously been made by Dr. Stefan Kerscher (called pKB86 and pKB87 by Dr. Katrine Bych)

### 2.5.7 Plasmid created and used in this study

**Table 2.8: List of plasmids**

Vector name	My reference	Source
Ind1-strep-pUB4	pAM3	Bych et al. 2008
Ind1-N271QfsX31-strep-pUB4	pAM4	Wydro et al. 2013
pUB26	pAM5	Ulrich Brandt
Ind1-L102P-strep-pUB4	pAM8	Created in this study
Ind1-G285C-strep-pUB4	pAM9	Created in this study
Ind1-L191F-strep-pUB4	pAM12	Created in this study
Ind1-L191F-strep-pUB4	pAM13	Created in this study
Ind1-G136D-strep-pUB4	pAM24	Created in this study
Ind1-M180V-strep-pUB4	pAM25	Created in this study
Ind1-M180A-strep-pUB4	pAM26	Created in this study
Δ36-Ind1-G80S-strep-pET15b	pAM22	Stefan Kerscher
Δ36-Ind1-K81Q-strep-pET15b	pAM23	Stefan Kerscher
Δ36-Ind1-strep-pET15b	pAM32	Bych et al. 2008
Δ36-Ind1-L102P-strep-pET15b	pAM33	Created in this study
Δ36-Ind1-D103Y-strep-pET15b	pAM34	Created in this study

## 2.6 Genetic transformation

### 2.6.1 *Escherichia coli* transformation

Aliquots of competent *Escherichia coli* cells (DH5α or TOP10) were thawed on ice. 50-100 ng plasmid was added, mixed, and incubated on ice for 30 minutes. Cells were then heat-shocked at 42°C for 45 seconds, before being put on ice for a further 2 minutes. 0.5 ml SOC medium (for 1 L: 20 g tryptone, 5 g yeast extract, 0.5 g NaCl pH 7.0, 0.186 g KCL, 20 mM glucose, 2 mM MgCl<sub>2</sub>) was added, and the mixture incubated at 37°C for 1 hour. Cells were then collected by centrifugation, resuspended in 100 μl SOC, and

spread on selective LB agar plates (for 1 L: 10 g tryptone, 5 g yeast extract, 10 g NaCl pH 7.0, 15 g agar) Transformants were visible after 1 day of growth at 37°C

### 2.6.2 *Yarrowia lipolytica* transformation

*Yarrowia lipolytica* cells were transformed as described in (Chen et al. 1997). Briefly, cells were grown overnight in Y $\frac{1}{2}$ D, at 28°C, and collected by centrifugation, or alternatively harvested from a fresh Y $\frac{1}{2}$ D plate. Cells were resuspended in 100  $\mu$ l one-step buffer (45% PEG4000, 0.1 M lithium acetate pH 6.0, 0.1M dithiothreitol, 25  $\mu$ g/100  $\mu$ l single-stranded carrier DNA) and 200-500 ng plasmid was added. The mixture was vortexed and then incubated at 39°C for 1 hour. The mixture was spread on a Y $\frac{1}{2}$ D plate containing 75  $\mu$ g ml $^{-1}$  Hygromycin B. Transformants were visible after 2-3 days growth at 28°C.

### 2.6.3 Antibiotics

**Table 2.9: List of antibiotics**

Antibiotic	Concentration ( $\mu$ g ml $^{-1}$ )	Solvent
Carbenicillin	100	H $_2$ O
Chloramphenicol	34	EtOH
Hygromycin	75	H $_2$ O
Tetracyclin	5	70% EtOH

## 2.7 *Arabidopsis thaliana* growth and manipulation

### 2.7.1 Growth on MS plates

*Arabidopsis thaliana* seeds were vernalised for 24 hours at 4°C and then were sterilised by incubation for 2 x 2 hours in chlorine gas (created by adding 100 ml bleach to 6 ml 5 M HCL). Seeds were grown on  $\frac{1}{2}$ MS agar plates ( $\frac{1}{2}$  MS salts, 0.8% (w/v) agar, 0.05% (w/v) MES-KOH, pH 5.7). When mutants with severe growth phenotypes were grown, 1% (w/v) sucrose was added to the plates to aid germination. Growth conditions were 16 hours light, 8 hours dark, 22°C, at light intensity 120-160  $\mu$ mol photons m $^{-2}$  s $^{-1}$ .

### 2.7.2 Growth on soil

*Arabidopsis* seeds were planted directly on soil after vernalisation for 24 hours at 4°C. Growth conditions were 16 hours light, 8 hours dark, 22°C at light intensity 180-200  $\mu$ mol photons m $^{-2}$  s $^{-1}$ .

### 2.7.3 Growth in hydroponic culture

After sterilisation *Arabidopsis thaliana* seeds were planted in a 250 ml flask in 50 ml hydroponic growth medium ( $\frac{1}{2}$  MS salts, 1% (w/v) sucrose 0.015% (w/v) agar, 0.05% (w/v) MES-KOH, pH 5.8). These were left stationary for 5 days, followed by 7 day shaking at 70 rpm. Growth conditions were 16 hours light, 8 hours dark, 22°C, at light intensity 120-160  $\mu\text{mol photons m}^{-2} \text{ s}^{-1}$ . 12 days after planting mitochondrial purification was performed (see section 2.91).

### 2.7.4 Callus cell lines

*Arabidopsis thaliana* were grown on vertical  $\frac{1}{2}$ MS agar plates until the roots were 2-3 cm in length. These were then cut off and placed on callus induction agar plates (0.386 % (w/v) Gamborg B5 medium (incl. vitamins), 2% (w/v) glucose, 0.05% (w/v) MES-KOH, pH 5.7, 0.8% (w/v) agar, 0.5 mg/L of 2,4-dichlorophenoxyacetic acid, 0.05 mg/L kinetin) and left for 3-4 weeks at 22°C. Callus tissue was then transferred to fresh callus induction plates, and sub-cultured every two weeks.

### 2.7.5 Cross pollination

*Arabidopsis thaliana* crosses were carried out as described on the Arabidopsis Stock Centre website ([http://arabidopsis.info/Info Pages?template= crossing;web\\_section= arabidopsis](http://arabidopsis.info/Info_Pages?template= crossing;web_section= arabidopsis)).

## 2.8 *Yarrowia lipolytica* growth

### 2.8.1 Agar plates

*Yarrowia lipolytica* cells were grown on Y $\frac{1}{2}$ D agar plates (1% (w/v) yeast extract, 1% (w/v) glucose, 2% (w/v) agar) at 30°C. When *Yarrowia lipolytica* cells contained a plasmid, 75  $\mu\text{g ml}^{-1}$  Hygromycin B was added for selection.

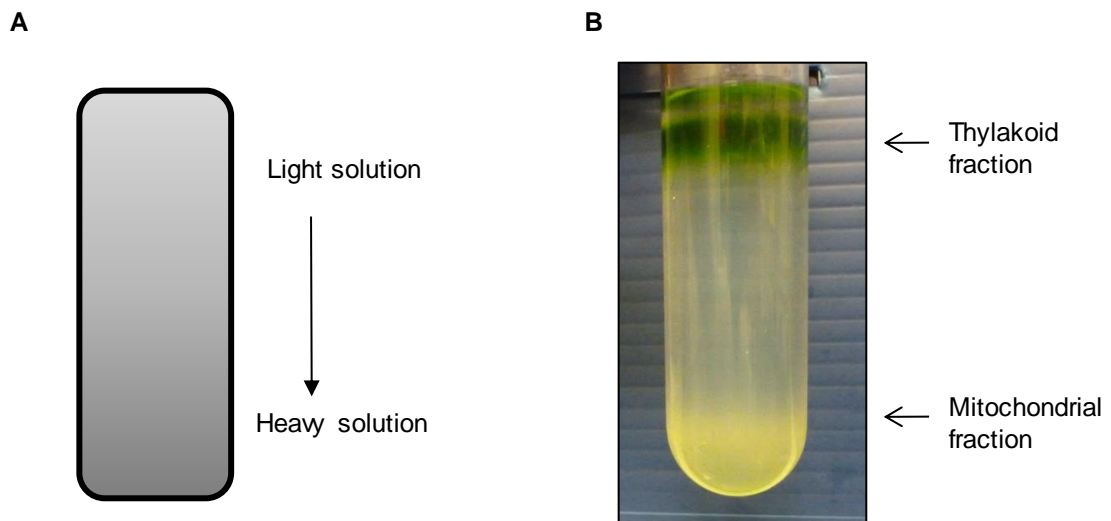
### 2.8.2 Liquid culture

*Yarrowia lipolytica* cells were grown in Y $\frac{1}{2}$ D liquid cultures (1% (w/v) yeast extract, 1% (w/v) glucose) at 28°C at 200 rpm. When *Yarrowia lipolytica* cells contained a plasmid, 75  $\mu\text{g ml}^{-1}$  Hygromycin B was added for selection.

## 2.9 Mitochondrial purification

### 2.9.1 Mitochondrial purification from *Arabidopsis thaliana* leaves or hydroponic culture

Mitochondria were isolated from *Arabidopsis thaliana* leaves or hydroponically grown seedlings as described in (Sweetlove et al. 2007) with minor modifications. Briefly, plant material was homogenised using a 2 cm Polytron® blender in 3 volumes of ice cold extraction buffer (0.3 M sucrose, 25 mM  $\text{Na}_2\text{P}_2\text{O}_7\text{NaOH}$ , pH 7.5, 2 mM EDTA, 10 mM  $\text{KH}_2\text{PO}_4$ , 0.5% (w/v) PVP-40, 0.2% (w/v) BSA, fresh 4 mM cysteine). The samples were then filtered through 2 layers of muslin and 1 layer of miracloth (Millipore). Differential centrifugation was performed at 2,000 x *g* for 5 minutes and 12,500 x *g* for 20 minutes. Pellets were resuspended in ice cold wash buffer (0.3 M sucrose, 10 mM TES-KOH, pH 7.5), before another round of differential centrifugation. Samples were then layered on continuous Percoll density gradients (see next paragraph; Figure 2.1A) and centrifuged at 40,000 x *g* for 40 minutes, with the breaks set on low. The mitochondria fraction (Figure 2.1B) was then collected and washed twice with wash buffer and centrifugation at 16,000



**Figure 2.1: Mitochondrial purification using a Percoll gradient**

**(A)** Schematic of Percoll gradient, showing distribution of light and heavy solution

**(B)** Percoll gradient after centrifugation. *Arabidopsis thaliana* sample has been separated into chloroplast thylakoid and mitochondrial fractions.

x g for 20 minutes. Further wash steps can be performed with a table top microcentrifuge as necessary. Protein concentration was then determined (see section 2.10.2).

Percoll gradients were made by adding 0.2% BSA (w/v) to 2 x wash buffer (to give 2 x gradient buffer) before adding Percoll to a final concentration of 28% (w/v) (for 150 mL: 75 ml 2 x gradient buffer, 42 ml Percoll and 33 ml dH<sub>2</sub>O). This makes the “light solution”. Half of this was taken and PVP-40 was added, to a concentration of 4.4% (w/v), which was then renamed the “heavy solution”. 16 ml each of the “heavy” and “light” solution were mixed with a gradient former (BioRad, model 385) to form a gradient. With some purifications, a 5 ml Percoll pad was made below the gradient (for 5 ml: 2 ml Percoll, 2.5 ml 2 x wash buffer, 0.5 ml dH<sub>2</sub>O). Mitochondria were stored at -80°C.

### **2.9.2 Mitochondrial purification from *Arabidopsis thaliana* callus**

Mitochondria were isolated from *Arabidopsis thaliana* callus culture using a protocol based on (León et al. 2007), adapted by Janneke Balk and Pia Sharma, 2008. Briefly, callus material was homogenised with a pestle and mortar with glass beads and 2 volumes ice cold extraction buffer (0.03 M MOPS-KOH, pH 7.5, 0.3 M mannitol, 1 mM EDTA, 0.1% (w/v) BSA, 0.6% (w/v) PVP-40, 4 mM cysteine). The samples were then filtered through 2 layers of miracloth (Millipore). Differential centrifugation was performed for 20 minutes at 1,000 x g and 20 minutes at 12,500 x g. Pellets were resuspended in ice cold wash buffer (0.03 M MOPS-KOH, pH 7.5, 0.3 M mannitol, 1 mM EDTA) before another round of differential centrifugation. Pellets were resuspended in wash buffer and protein concentration was determined (see section 2.10.2). Mitochondria were stored at -80°C.

### **2.9.3 *Yarrowia lipolytica* mitochondrial membranes**

Unsealed mitochondrial membranes were isolated based on the protocol in (Kerscher et al. 1999), with minor modifications. Briefly, *Yarrowia lipolytica* cells were grown in a 25 ml starter culture, Y<sub>1/2</sub>D containing 75 µg ml<sup>-1</sup> Hygromycin B, at 28°C, two days before the mitochondrial purification. The day before, the 25 ml starter culture was added to a 0.5 L culture, and grown overnight at 28°C. Cells were harvested by centrifugation at 3,500 x g, after growth had reached log stage. Pellet weights were typically between 2-8 g. Cells were washed with dH<sub>2</sub>O. To X g cells was added 2X g fine glass beads, 2 mM PMSF and 2X ml ice cold mito-membrane buffer (0.6 M sucrose, 20 mM MOPS-NaOH, pH7.5, 1 mM EDTA). Cells were disrupted by 15 rounds of 1 minute vortexing and 1 minute incubation in ice. Differential centrifugation was performed at 3,500 x g for 10



minutes and 40,000 x *g* for 120 minutes. Pellets were resuspended in mito-membrane buffer and stored at -80°C.

## 2.10 SDS-PAGE and Western blotting

### 2.10.1 Total protein isolation *Arabidopsis thaliana* and *Viscum album*

Leaf, or bud, tissue was frozen in liquid nitrogen and pulverised with a mini-pestle in a 1.5 mL tube. Lysis buffer (50 mM Tris-HCL pH 8.5, 5% glycerol (v/v), 1% SDS (w/v), 10 mM EDTA, 1 mM PMSF) was added, and mixed in. Debris was removed by centrifugation at 16,100 x *g* for 10 minutes at 4°C, and the supernatant taken for SDS-PAGE.

### 2.10.2 Protein determination

Protein concentration was determined using Bio-Rad Protein Assay Dye Reagent. 1-2 µl of sample was added to dH<sub>2</sub>O, to make up to 100 µl. 1 mL of dye (diluted 1 in 5 from concentrate) was added. Absorption at 595 nm was measured using a Thermo Spectronic spectrophotometer. Results were compared to a calibration curve, made at the same time, of Bovine Serum Albumin (BSA), ranging from 2-16 µg/µL.

### 2.10.3 SDS-PAGE

SDS-PAGE was performed according to (Laemmli 1970). Briefly, SDS-PAGE gels were made according to table 2.10, using a mini-gel system (Bio-Rad) with 0.75 mm spacers. Gel acrylamide percentages were set to 12.5 or 15%, as appropriate for optimal target protein separation. Samples were mixed with 4 x sample buffer (0.5 M Tris-HCL, pH 6.8, 8% (w/v) SDS, 40% (v/v) glycerol), to give a final concentration of 1 X buffer. Samples were typically between 5-50 µg/ lane, depending on sample type and target protein. Gels were run at ~10 mA through the stacking gel, and 20-30 mA through the separating gel, until the dye front exited the base of the gel.

**Table 2.10: SDS-PAGE gel composition**

Solution	Composition
Separating gel (for 15%)	2.5 mL 30% acrylamide 0.8% bis-acrylamide, 1.25 mL 1.5 M Tris-HCL, pH8.8, 0.4% SDS, 1.25 mL dH <sub>2</sub> O, 50 µL 10% APS, 5 µL TEMED.
Stacking gel	0.375 mL 30% Acrylamide 0.8% Bis-Acrylamide, 0.625 mL 0.5 M Tris-HCL, pH6.8, 0.4% SDS, 1.5 mL dH <sub>2</sub> O, 25 µL 10% APS, 2.5 µL TEMED.

### 2.10.4 Immunoblotting and ECL detection

Protein was transferred from the SDS-PAGE gel to a nitrocellulose membrane (0.2  $\mu\text{m}$ , Protran™) using a semi-dry blotting chamber (Biometra, Analytik Jena) and transfer buffer (0.048 M Tris, 0.039 M glycine, 0.0375 (w/v) SDS, 20% (v/v) methanol), for 30-35 minutes at 175-190 mA, as described (Towbin et al. 1979). After protein visualisation (see section 2.10.5), membranes were blocked for 1-12 hours with Tris-Buffered Saline (TBS) solution (50 mM Tris-HCL pH 7.5, 150 mM NaCl), containing 5% (w/v) milk powder and 0.1% (v/v) Tween-20 (TBS-TM). The membrane was then incubated in primary antibody for 2 hours. Antibodies raised in rabbits were made up to the correct dilution in TBS-TM. Antibodies raised in mice were made up to the correct dilution in Phosphate-Buffered Saline (PBS) solution, containing 1% BSA. Residual primary antibody was washed with 3 x 10 minute washes, with TBS-TM or PBS, as appropriate. Membranes were then incubated with the appropriate secondary antibody horseradish peroxidase conjugate (Goat anti-rabbit IgG, H&L, HRP or Goat anti-mouse IgG, H&L, HRP) made up in TBS-TM or PBS as appropriate, for 45 minutes. After final washing, membranes were incubated with Enhanced ChemiLuminescence (ECL) reagents (1:1 mixture of solution 1 and solution 2, see table 2.11) for 1 minute. Membranes were dried with filter paper before detection with a ImageQuant LAS500 (GE Healthcare) or high performance chemiluminescence film (Amersham Hyperfilm™ ECL) and x-ray film processor machine (Konica Minolta SRX-101A). Numerous exposures were taken, for different lengths of time.

**Table 2.11: ECL reagents**

Solution	Composition
Solution 1	0.1 M Tris-HCL, pH 8.0, 10 $\mu\text{l m}^{-1}$ luminol stock, 4.4 $\mu\text{l ml}^{-1}$ coumaric acid stock
Solution 2	0.1 M Tris-HCL, pH 8.0, 0.018% $\text{H}_2\text{O}_2$
Luminol stock	4.4% 3-aminophthalhydrazine (w/v) in DMSO
Coumaric acid stock	1.5% p-coumaric acid (w/v) in DMSO

### 2.10.5 Ponceau and InstantBlue™ stain

Total protein was stained to control for equal loading and transfer using either Ponceau S stain or InstantBlue™ (Expedeon). For Ponceau, after transfer the nitrocellulose membrane was incubated with Ponceau S solution (0.1% Ponceau S (w/v), 5% acetic acid (v/v)), before being washed off with  $\text{dH}_2\text{O}$ . Pictures were then taken before

continuation of immunoblotting. For in gel protein staining, the SDS-PAGE gel was incubated the SDS-PAGE gel was incubated for ~12 hours in InstantBlue™ (Expedeon), before being dried in and imaged.

## 2.10.6 Antibodies

**Table 2.12: List of antibodies**

Protein	Organism	Rabbit (R) or Mouse (M)	Dilution	Apparent size (kDa) (approximate)	Source
Aconitase	<i>S. cerevisiae</i>	R	1:2000	80	R. Lill
Aconitase 1	<i>A. thaliana</i>	R	1:5000	98	Agrisera
Cox2	<i>A. thaliana</i>	R	1:2000	30	Agrisera
E1 $\alpha$ (PDH)	<i>Z. mays</i>	M	1:1000	45	C. Leaver, T. Elthon
GRXS15	<i>A. thaliana</i>	R	1:1000	15	N. Rouhier
H-protein	<i>A. thaliana</i>	R	1:1000	Various	O. Keech
Ind1	<i>Y. lipolytica</i>	R	1:2000	30	R. Lill
INDH	<i>A. thaliana</i>	R	1:2000	30	R. Lill
ISU1	<i>A. thaliana</i>	R	1:5000	15	S. Lobreaux
Lipoate	N/A	R	1:3000	Various	Abcam- ab58724
Nad1	<i>A. thaliana</i>	R	1:5000	N/A	E. Meyer
Nad9	Wheat	R	1:5000	23	N/A
NBP35	<i>A. thaliana</i>	R	1:2000	38	R. Lill
NFU4	<i>A. thaliana</i>	R	1:5000	Various	N. Rouhier
NUBM	<i>Y. lipolytica</i>	M	1:20	50	V. Zickermann
NUCM	<i>Y. lipolytica</i>	M	1:20	50	V. Zickermann
Porin	<i>Z. mays</i>	M	1:1000	32	C. Leaver, T. Elthon
Sdh2	<i>S. cerevisiae</i>	R	1:5000	28	R. Lill
TOM40	<i>A. thaliana</i>	R	1:5000	36	J. Whelan

## 2.11 Blue Native PAGE

### 2.11.1 *Arabidopsis thaliana* and *Viscum album*

To make samples, 100  $\mu$ g pellet *Arabidopsis thaliana* or *Viscum album* mitochondria was resuspended in 90  $\mu$ l ACA750 (750 mM ACA, 50 mM Bis-Tris-HCL, pH 7.0, 0.5 mM

EDTA). 10 µl of 10 % (w/v) n-dodecylmaltoside (DDM) in dH<sub>2</sub>O was added, mixed, and incubated on ice for 5 minutes followed by centrifugation at 16,100 x g, at 4°C, for 10 minutes. 95 µl supernatant was then added to 5 µl 5% Serva blue solution (ACA750 with 5% (w/v) Coomassie G250). 25 µl (25 µg) was loaded per lane. 5 µl NativeMark™ was used as a molecular weight marker. All BN-PAGE performed used NativePage™ 4-16% Bis-Tris gels (Invitrogen), with 1 mm spacers. Anode buffer was 50 mM Bis-Tris-HCL, pH7.0, cathode buffer was 50 mM Tricine, 15 mM Bis-Tris-HCL, pH7.0. To the cathode buffer 0.02% Coomassie G250 was added to make “dark” solution. This was diluted 10 times to make “light” solution (for 1 L: 100 ml dark solution, 200 ml 5 x cathode buffer, 700 ml dH<sub>2</sub>O). Gels were run in a XCell SureLock™ chamber at 4°C at 100 V (max 10 mA) for 45 minutes with “dark” solution”, before being exchanged for “light” solution and being run for then ~2.5 hours at 250 V (max 15 mA), until the dye front exited the bottom.

### **2.11.2 *Yarrowia lipolytica***

To make samples, unsealed mitochondrial membranes were mixed with solubilisation buffer (90% of ACA750, 10% of 10 % (w/v) DDM) and incubated on ice for 5 minutes followed by centrifugation at 16,100 x g, at 4°C, for 10 minutes. A Bradford assay was carried out on the supernatant (see section 2.10.2), and 25 µg taken. This was then made up to 95 µl with solubilisation buffer and added to 5 µl 5% Serva blue solution. 25 µl (6.25 µg) was loaded per lane. BN-PAGE was then carried out as described in section 2.11.1.

### **2.11.3 Transfer to PVDF membrane and immunolabelling**

Protein was transferred from the BN-PAGE gel to a PVDF membrane (0.2 micron, Millipore™) using wet, or tank, transfer (Towbin et al. 1979) and BN cathode buffer, for 600 minutes at 40 mA, (maximum voltage 100 V). After transfer the membrane can be used for Coomassie staining (see section 2.11.4), complex III direct detection (see section 2.11.6) or immunolabelling (see section 2.10.4).

### **2.11.4 Coomassie stain**

BN gels or PVDF membrane were stained in Coomassie solution (45% (v/v) MeOH and 5% (v/v) acetic acid, 0.05% Coomassie R250) for 1-12 hours. Gels were destained using a solution of 45% (v/v) MeOH and 5% (v/v) acetic acid.

### **2.11.5 Complex I, II and IV activity stain**

BN gels were stained for complex I, II and IV activity as described in (Sabar et al. 2005). Briefly, BN gels were equilibrated in buffer for 10 minutes before incubation in staining

solution (Table 2.13). After desired staining intensity was reached (20 minutes for complex I, up to 2 hours for complex IV), the reaction was stopped using a solution of 45% (v/v) MeOH and 5% (v/v) acetic acid. Destaining using a solution of 45% (v/v) MeOH and 5% (v/v) acetic acid was performed for optimal enzyme stain visualisation.

**Table 2.13: BN-PAGE in gel activity stains**

Respiratory complex and enzyme activity	Buffers	pH	Substrates	Reagents (%(w/v))
Complex I (NADH dehydrogenase)	0.1 M Tris HCL	7.4	0.2 mM NADH	0.1 NBT
Complex II (succinate dehydrogenase)	50 mM KH <sub>2</sub> PO <sub>4</sub> , 0.1 mM ATP, 0.2 mM PMS	7.4	10 mM succinate	0.2 NBT
Complex IV (cytochrome c oxidase)	50 mM KH <sub>2</sub> PO <sub>4</sub>	7.4	1 mg ml <sup>-1</sup> cytochrome c	0.1 DAB

### 2.11.6 Complex III detection

Complex III can be detected by ECL (see section 2.10.14), immediately after transfer to a PVDF membrane (see section 2.11.3), as described (Feissner et al. 2003; Sun et al. 2015).

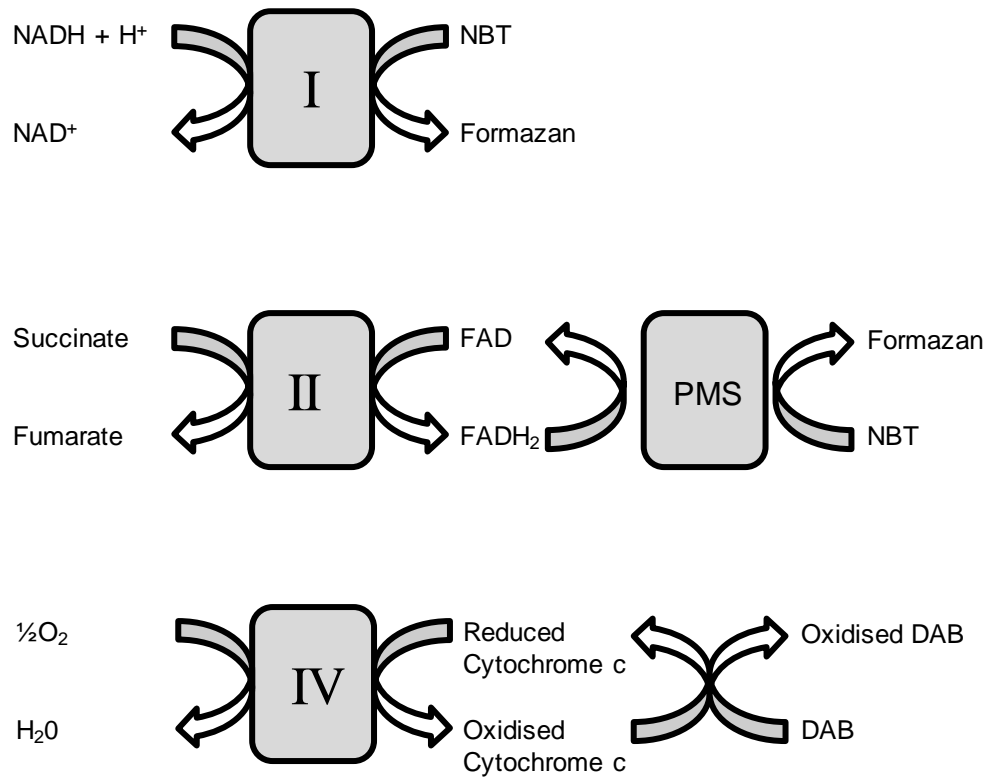
## 2.12 Spectrophotometric enzyme assays

### 2.12.1 Complex I NADH:HAR oxidoreductase assays

NADH:HAR oxidoreductase activity was assayed as described in (Abdrakhmanova et al. 2005) and (Tocilescu et al. 2007). Briefly, NADH:HAR oxidoreductase activity was measured as NADH oxidation ( $\epsilon_{340-400\text{ nm}} = 6.22\text{ mM}^{-1}\text{ cm}^{-1}$ ) in the presence of the artificial electron acceptor hexaammineruthenium(III) chloride (HAR). A 1 ml, 10 mm, quartz cuvette (Hellma® Analytix) was used and measurements made on a UV/Vis spectrophotometer (Jasco, V-550). Activity for 25 µg mitochondria was measured in 20 mM HEPES-NaOH, pH 8.0, 2 mM NaN<sub>3</sub>, 0.2 mM NADH, 2 mM HAR and 40 µg ml<sup>-1</sup> alamethicin. Measurement started after addition of HAR.

### 2.12.2 Complex I dNADH:DBQ oxidoreductase assay

dNADH:DBQ oxidoreductase activity was assayed as described in (Abdrakhmanova et al. 2005) and (Tocilescu et al. 2007). Briefly, dNADH:DBQ oxidoreductase activity was



**Figure 2.2: Histochemical stains of the mitochondrial respiratory chain**

Schematic showing histochemical stain activities. Adapted from Sabar et. al., 2005. Nitro-blue tetrazolium (NBT) gets reduced to formazan by the action of complex I. NBT gets reduced to formazan by the action of complex II, along with the artificial electron carrier phenazine methosulphate (PMS). 3,3'-Diaminobenzidine (DAB) is oxidised by cytochrome c, which is connected to the action of complex IV.

measured as dNADH oxidation activity ( $\epsilon_{340-400\text{ nm}} = 6.22\text{ mM}^{-1}\text{ cm}^{-1}$ ) in the presence of the ubiquinone analogue *n*-decylubiquinone (DBQ) as an electron acceptor. A 1 ml, 10 mm, quartz cuvette (Hellma® Analytix) was used and measurements made on a UV/Vis spectrophotometer (Jasco, V-550). Activity for 50 µg mitochondria was measured in 20 mM MOPS-NaOH, pH 7.4, 50 mM NaCl, 2 mM KCN, 0.1 mM dNADH and 0.1 mM DBQ. Measurement started after addition of DBQ. 0.01 mM of the complex I inhibitor Piericidin A was added at the end to ensure that dNADH:DBQ oxidoreductase activity is due to complex I.

## 2.13 Microscopy

### 2.13.1 Light microscopy

Light microscopy for *Arabidopsis thaliana* silique dissection was carried out on a Leica M205FA light microscope.

### 2.13.2 Differential Interference Contrast

*Arabidopsis thaliana* embryos were prepared and imaged as described in (Andriotis et al. 2010), with help from Dr. Vasilios Andriotis. Briefly, seeds were cleared overnight in Hoyer's solution (Liu & Meinke 1998), without gum Arabic. Cleared seeds were then viewed under Differential Interference Contrast (DIC) optics on a Leica DM600.

## 2.14 Proteomics

### 2.14.1 Sample preparation and mass spectrometry

Mitochondrial samples from *Viscum album* were digested with trypsin. Briefly, samples were dissolved in a solution of 6 M urea, 2 M thiourea, pH 8.0. To this was added 1 µL reduction buffer (6.5 mM DTT) per 50 µg protein and the mixture was incubated for 30 minutes at room temperature. 1 µL alkylation buffer (27 mM iodoacetamide) per 50 µg protein was added and the mixture was incubated for 20 minutes at room temperature in the dark. 0.5 µL LysC (0.5 µg/µL) per 50 µg protein was added and incubated for 3 hours at room temperature. Samples were diluted 4X with 10 mM Tris-HCL, pH8.0, and to this 1 µL trypsin per 50 µg protein was added and incubated overnight at 37 °C. Next, the samples were acidified to a final concentration of 0.2% trifluoroacetic acid. Samples were then purified using ZipTip® (Promega) tips. Shotgun proteomics was then performed using a Thermo Q-Exactive Plus mass spectrometer, performed with help from Dr. Etienne Meyer.

### 2.14.2 Analysis

The output from mass spectrometry was analysed using the MASCOT server, using the following settings: enzyme: trypsin, allow up to 1 missed cleavage; fixed modification: carbamidomethyl; variable modification: methionine oxidation; peptide tolerance: 10 ppm; MS/MS tolerance: 0.6 Da; peptide charge: 2+, 3+ and 4+. Two databases were used to identify peptide fragments. The first was based on the *Arabidopsis thaliana* TAIR 10 gene annotation, with modifications made by Dr. Etienne Meyer to account for contaminants and edited organellar proteins. The second was based on the mitochondrial genome of *Viscum album*, sequenced by Petersen et al. 2015 (GenBank KJ129610), reannotated by Skippington et al. 2017.

### 2.15 Protein expression

Ind1-strep was expressed and purified as described in (Bych et al. 2008), using plasmids from section (table 2.8). Briefly, plasmids were transformed into *Escherichia coli* Rosetta pISC pLYS. Colonies were grown overnight at 37 °C in LB media, with correct antibiotics, then diluted 50 x in 100 ml Terrific broth (for 100 ml:4.76 g Terrific broth, 0.8 ml 50% (v/v) glycerol), and grown until OD<sub>600</sub>= 0.6. Protein expression was induced by addition of 1 ml L<sup>-1</sup> benzylalcohol, 50 µM Fe-ammonium citrate, 100 µM L-cysteine and 1.2 ml L<sup>-1</sup> IPTG. Cells were then grown at 20 °C overnight. Cells were collected by centrifugation at 5000 x g and flash frozen in liquid nitrogen. Cells were then resuspended in buffer W (100 mM Tris-HCL, pH 8.0, 150 mM NaCl) with 0.2% DDM. Cells were then sonicated, and separated into soluble and insoluble fractions by centrifugation at 16,100 x g. Western blot analysis was performed to confirm protein expression.



### **3 Using *Yarrowia lipolytica* as a model to understand pathogenic mutations in *NUBPL* associated with complex I deficiency**

#### **3.1 Introduction**

Defects in the mitochondrial respiratory chain underlie a large proportion of mitochondrial disorders (Calvo & Mootha 2010). Respiratory chain disorders are serious genetic diseases with an estimated prevalence of 1 in 5000 live births (Skladal et al. 2003) and 1 in 4300 in the adult population (Gorman et al. 2015) making them the most common inborn error of metabolism (Calvo & Mootha 2010). Complex I deficiency is the most common respiratory chain defect (Fassone & Rahman 2012; Mimaki et al. 2012; Rodenburg 2016) and is estimated to make up between a quarter and a third of all cases (Loeffen et al. 2000; Bugiani et al. 2004; Fassone & Rahman 2012). This may be because, with the largest number of subunits, complex I may be the most likely to be affected by genetic mutations (Rodenburg 2016). Clinical presentation usually occurs in infancy or early adulthood. Symptoms include Leigh syndrome, skeletal muscle myopathy, cardiomyopathy, hypotonia, stroke, ataxia and lactic acidosis (Fassone & Rahman 2012; Koene et al. 2012; Nouws et al. 2012; Alston et al. 2017). Such clinical heterogeneity, along with the varied genetic causes, make diagnosis challenging.

Of the 44 structural subunits, 37 are encoded in the nuclear genome and 7 in the mitochondrial genome. Disease causing mutations have been identified in 26 out of 44 subunit-encoding genes (Rodenburg 2016; Alston et al. 2017; Craven et al. 2017). However, this only provides a diagnosis for about 50% of cases (Calvo & Mootha 2010; Nouws et al. 2012). The remaining 50% are made up of other mutations, some of which affect other mitochondrial processes, such as mitochondrial DNA metabolisms, and some of which are complex I assembly factors.

Assembly factors are defined as proteins which are required for the correct assembly and function of complex I, but are not present in the mature structure. The first assembly factors were discovered in the model yeast *Neurospora crassa* in 1998 (Küffner et al. 1998), but since then numerous others have been discovered in humans (Mimaki et al. 2012; Nouws et al. 2012). Disease causing mutations have been discovered in 10 out of 16 currently known assembly proteins (Alston et al. 2017; Formosa et al. 2017). Discovery of these mutations has been greatly accelerated by advances in next

generation sequencing technologies, including whole exome sequencing (WES) technologies, which make it practical and affordable to sequence whole patient exomes (Haack et al. 2010; Alston et al. 2017). This has led to a greater understanding of the causes of mitochondrial disorders and an overall improvement in genetic diagnosis of patients.

One assembly factor of complex I is NUBPL (NUcleotide Binding Protein-Like). It was first identified in the aerobic yeast *Yarrowia lipolytica* where it is called Ind1 (Iron sulfur protein required for NADH dehydrogenase) (Bych et al. 2008). *Yarrowia lipolytica* has been used extensively in biotechnology but it has also been developed as model organism to study complex I biology (Kerscher et al. 2002; Kerscher et al. 2004). This has been necessary because the usual yeast model Baker's yeast, *Saccharomyces cerevisiae*, has lost complex I. A genomic deletion strain of *IND1* displayed a major decrease in fully assembled complex I, to approximately 28% of wild type (Bych et al. 2008). Given Ind1's homology to NBP35 and CFD1, cytosolic scaffold proteins involved in FeS cluster assembly, and its ability to coordinate a labile [4Fe-4S] cluster *in vitro*, it was suggested to be responsible for inserting FeS clusters of Complex I. Phylogenetic analysis of *IND1* showed that it is present in almost all eukaryotes, and closely matches the distribution of Complex I (Bych et al. 2008). The *IND1* homologue (*INDH*) in the model plant *Arabidopsis thaliana* is also required for complex I assembly (Wydro et al. 2013). Studies investigating the human homologue of *IND1*, initially called *huIND1* but now the official name is *NUBPL*, in HeLa cells showed that siRNA knockdown lead to substantial decreases in complex I and its subunit levels, as well as accumulation of complex I membrane arm assembly intermediates (Sheftel et al. 2009). Following that, mutations in *NUBPL* were found associated with complex I deficiency (Calvo et al. 2010; Tenisch et al. 2012; Kevelam et al. 2013). Recently a large scale mutant screen for respiratory defect phenotypes, based upon a mutant's inability to grow on galactose, identified *NUBPL* (Arroyo et al. 2016). Combined, these studies all indicate a role for NUBPL complex I assembly, confirming its conserved function across eukaryotes.

To date, 13 patients have been found with mutations in *NUBPL* (Table 3.1). Patients display various symptoms including mitochondrial encephalomyopathy and neurological defects. Magnetic resonance imaging (MRI) patterns of *NUBPL* patients are unique and distinct from other complex I deficiency patients and so can be used diagnostically (Kevelam et al. 2013). All patients to date have been compound heterozygous, where they inherit one copy of *NUBPL* with a branch-site mutation c.815-27T>C, and one copy of *NUBPL* with either a large deletion or a missense mutation resulting in an amino acid

exchange. The c.815-27T>C mutation affects a splice recognition site in the intron and leads to aberrant mRNA splicing (Tucker et al. 2012). Approximately 30% of wild-type *NUBPL* transcript remains but two mis-spliced transcripts are also produced. One of these disappears by nonsense mediated decay. The other, which skips exon 10 and fuses exon 9 and 11 as well as introducing a frameshift, is predicted to lead to a new protein product, D273QfsX31. A study in *Yarrowia lipolytica* where the mutant protein was recreated in the yeast homolog Ind1, as N271QfsX31, suggests that c.815-27T>C is pathogenic (Wydro & Balk 2013). Ind1 variant N271QfsX31 was unable to support wild-type levels of complex I assembly. Interestingly c.815-27T>C is often co-inherited with a missense c.G166A mutation (p.G56R), which, on its own, is not thought to be pathogenic (Tucker et al. 2012; Wydro & Balk 2013). As well as the c.815-27T>C mutation causing an estimated 70% decrease in wild-type *NUBPL* transcript, the remaining protein product D273QfsX31 is unstable, as *NUBPL* protein is almost undetectable in patient fibroblasts (Tucker et al. 2012). In *Yarrowia lipolytica* the Ind1 variant N271QfsX31 protein is also unstable and degrades (Wydro & Balk 2013). Therefore, if the mutation in the second allele affects a crucial amino acid and renders the protein non-functional, then complex I assembly would be severely impaired in the patient.

Studying patient mutations in a *Yarrowia lipolytica* model is useful for clinicians and clinical scientists. Complex I defects seen in patient tissue biopsies can often not be replicated in cultured fibroblasts making characterisation of patient's mutations hard. Data from WES show a large number of polymorphisms, many of which are benign variants, and it can be hard to pick the causal one. For example, a study claimed that 27% of suggested disease causing mutations were in fact polymorphisms, sequencing errors, or did not have sufficient evidence of pathogenicity (Bell et al. 2011). Therefore, functional characterisation of patient mutations is essential to assign pathogenicity. *Yarrowia lipolytica* is a model which is well characterised, amenable to genetic manipulation, easy and cheap to grow, and is therefore ideal for this task.

Here, in an effort to assign pathogenicity to six *NUBPL* mutations, four of which are newly identified and unpublished, I recreated the corresponding mutations in the *Yarrowia lipolytica* homolog, Ind1, and studied their effect. This approach has been previously used to model the effect of pathogenic mutations on complex I structural subunits, in both *Yarrowia lipolytica* and *Neurospora crassa* models (Ahlers et al. 2000; Kerscher et al. 2004; Duarte et al. 2005; Ferreira et al. 2011; Varghese et al. 2015; Gerber et al. 2016).

## 3.2 Results

### 3.2.1 Selection of *NUBPL* mutations for study

To date, 13 patients with mitochondrial disorders have been identified who have mutations in *NUBPL* (Table 3.1). In addition, many more mutations in *NUBPL* have been uncovered by large scale exome sequencing projects (exac.broadinstitute.org).

Mutations that result in frameshifts, large scale insertions or deletions (patients 2,5,7 and 10) are almost certainly deleterious and so were not selected for further study. Mutations that affect splicing would be hard to recreate in *Yarrowia lipolytica* *IND1*, which is a single exon gene. One exception to this is the branch-site mutation c.815-27T>C which results in a decreased amount of wild-type mRNA and a truncated protein product D273QfsX31. This was previously recreated in *Ind1*, by gene synthesis, resulting in *Ind1* N271QfsX31 (Wydro & Balk 2013) and found to be deleterious. It will be used as a control throughout.

Mutations resulting in a single amino acid change at a conserved position were chosen for study (Table 3.2): L104P (patient 8,9,12), D105Y (patient 3,4), G138D (ExAC database), L193F (patient 6) and G287C (patient 11). L104P and D105Y affect residues in the highly-conserved Switch I motif in the P-loop which is needed for NTP hydrolysis; G138D and L193F affect residues in the unstudied MRP domain and G287C introduces an additional cysteine to the C-terminus. G138D is not associated with a clinical case but is present in exome databases at high incidence (Table 3.4). G136D has also been suggested to be connected to Parkinson's disease (Dr. Eli Hatchwell and Dr. Peggy Eis, personal communication). Therefore, because of current research on the connection between mild complex I defects and Parkinson's disease, it was included in this study. In order to identify the corresponding amino acids in *Yarrowia lipolytica* *Ind1*, an alignment of amino acid sequences of *NUBPL* and *Ind1* was made using Clustal Omega (Figure 3.1, Table 3.2). A homology model of *Ind1* was made and positions of amino acids studied here marked (Figure 3.2). These protein variants will be studied in the remainder of this chapter in order to assess their effect on *Ind1* function.

### 3.2.2 Bioinformatic analysis of *NUBPL* variants

Before biochemical analysis, much can be learned from bioinformatic analysis. In order to predict the effect of *NUBPL* variants the prediction programme PolyPhen-2 was used.

**Table 3.1: Overview of *NUBPL* mutations associated with complex I deficiency**

This table combines both published and unpublished reports. The genotype is given as HM for homozygous, or HT for heterozygous. Inheritance is given as paternal or maternal, where known. Nomenclature according to Human Genome Variation Society.

Patient	Country of origin	cDNA	Protein	Site	Paternal or Maternal	Reference
1	Argentina	c.166G>A	p.G56R	Exon 2	N/A	Kevelam et al. 2013
		c.815-27T>C <sup>d</sup>	p.D273QfsX31	Intron 9	N/A	
2	Germany	c.166G>A	p.G56R	Exon 2	Paternal	Kevelam et al. 2013
		c.815-27T>C <sup>d</sup>	p.D273QfsX31	Intron 9	Paternal	
		c.667_668insCC TTGTGCTG	p.Q223Afs*4	Exon 8	Maternal	
3,4	Canada	c.166G>A	p.G56R	Exon 2	Paternal	Kevelam et al. 2013
		c.815-27T>C <sup>d</sup>	p.D273QfsX31	Intron 9	Paternal	
		c.313G>T	p.D105Y	Exon 4	Maternal	
5	USA	c.166G>A	p.G56R	Exon 2	Paternal	Kevelam et al. 2013
		c.815-27T>C <sup>d</sup>	p.D273QfsX31	Intron 9	Paternal	
		c.693+1G>A <sup>e</sup>	p.?	Intron 8	Maternal	
6	Netherlands	c.579A>C	p.L193F	Exon 7	Paternal	Kevelam et al. 2013
		c.166G>A	p.G56R	Exon 2	Maternal	
		c.815-27T>C <sup>d</sup>	p.D273QfsX31	Intron 9	Maternal	
7	Australia	c.166G>A	p.G56R	Exon 2	Paternal	Calvo et al 2010; Tucker et al 2012
		c.815-27T>C <sup>d</sup>	p.D273QfsX31	Intron 9	Paternal	
		240-kb deletion		Exons 1-4	Maternal	
		137-kb duplication		Exon 7	Maternal	
8,9	USA	c.815-27T>C	p.D273QfsX31	Intron 9	Paternal	Virginia Kimonis, pers. comm
		c.311T>C	p.L104P	Exon 4	Maternal	
10	France	c.815-27T>C	p.D273QfsX31	Intron 9	Paternal	Tenish et al. 2012
		c.205_206delGT	p.V69Yfs*80		Maternal	
11	Germany	c.166G>A	p.G56R	Exon 2	N/A	Holger Prokisch, pers. comm
		c.815-27T>C	p.D273QfsX31	Intron 9	N/A	
		c.859G>T	p.G287C	Exon 10	N/A	
12	USA	c.815-27T>C	p.D273QfsX31	Intron 9	N/A	Baylor Miraca Genetics Laboratories
		c.166G>A	p.G56R	Exon 2	N/A	
		c.311T>C	p.L104P	Exon 4	N/A	
13	USA	c.545T>C	p.V182A	Exon 7	N/A	Daryl Scott, pers. comm
		c.815-27T>C	p.D273QfsX31	Intron 9	Paternal	

**Table 3.2: Protein variants in *Homo sapiens* NUBPL and *Yarrowia lipolytica* Ind1**

NUBPL amino acid variant	Corresponding Ind1 amino acid variant
L104P	L102P
D105Y	D103Y
G138D	G136D
L193F	L191F
G287C	G285C
D273QfsX31	N271QfsX31

**Table 3.3: PolyPhen-2 predictions of effect of NUBPL variants**

NUBPL amino acid variant	Polyphen-2	
	Class	Score
L104P	Probably damaging	1.000
D105Y	Probably damaging	1.000
G138D	Probably damaging	0.999
V182A	Possible damaging	0.923
L193F	Probably damaging	1.000
G287C	Probably damaging	1.000

**Table 3.4: Allele frequency of mutations leading to NUBPL variants from ExAC**

Variant	Allele Count	Total Alleles	Allele Frequency
L104P	25	120170	$2.08 \times 10^{-4}$
D105Y	1	120200	$8.319 \times 10^{-6}$
G138D	88	117184	$7.51 \times 10^{-4}$
V182A	348	120522	$2.887 \times 10^{-3}$
c.815-27T>C	436	117040	$3.725 \times 10^{-3}$

### Chapter 3 Using *Yarrowia lipolytica* as a model to understand pathogenic mutations in NUBPL associated with complex I deficiency

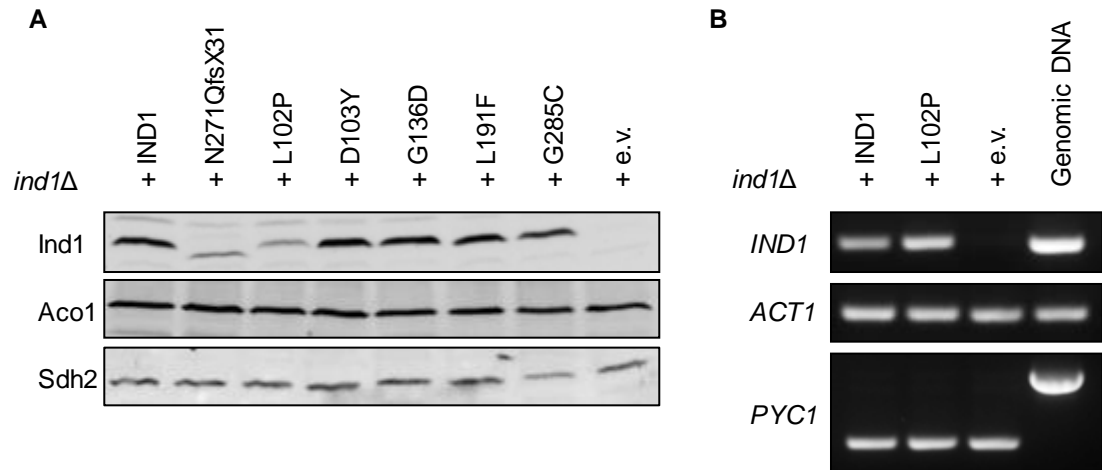
HsNUBPL	1	MGIWQRLLLF	GVSLRAGGGAT	API	GG	SRAMV	CGR	QL	-----	SGAG	SETL
YlInd1	1	-----MRGF	-----RLI	API	QR	STAI	ISRL	Q	PITANFHSSPALR	SHENPL	GIPKS

PolyPhen-2 predicts whether mutations will be deleterious or not based upon sequence conservation (Adzhubei et al. 2010). Analysis of *NUBPL* protein variants by Polyphen-2 returns a result of “probably damaging” for all of them (Table 3.3). Therefore, the base assumption before performing experiments is that all mutations will be deleterious.

Additional insight can be gained from a population genetics approach. The prevalence of *NUBPL* mutations in populations can hint at their severity. Deleterious mutations are usually present at lower levels than less deleterious, or benign, mutations. Alternatively, high levels of a mutation can be due to positive selection. In order to assess the severity of *NUBPL* mutations, prevalence was assessed using data from ExAC (exac.broadinstitute.org), which provides global and population level data (Table 3.4). Variant D105Y had the lowest allele frequency, by two orders of magnitude, followed by L104P and G138D. This suggests that these three mutations are the most highly deleterious. The highest allele frequency was for the branch site mutation, c.815-27T>C, which was present at a frequency of  $3.725 \times 10^{-3}$ , a remarkably high prevalence for a potentially deleterious mutation. Three individuals were identified which were homozygous for this variant. However, no clinical data exists on these individuals.

The low global prevalences of these mutations hide regional differences. For example, G138D has a higher allele frequency in non-Finnish Europeans ( $1.152 \times 10^{-3}$ ) than globally ( $7.510 \times 10^{-4}$ ). The biggest difference is found with c.815-27T>C which has allele frequencies in European populations (non-Finnish) of  $4.945 \times 10^{-3}$  and in Finnish Europeans of  $1.255 \times 10^{-2}$ . This last figure would imply around 1% of the Finnish population have this allele. As (Tucker et al. 2012) pointed out, such prevalence would make it the most common autosomal recessive respiratory chain mutation. Indeed, based upon data from HapMap and the 1000 genomes project, (Tucker et al. 2012) calculated that approximately 1 in 14,000 individuals would be heterozygous for this mutation. If such individuals had OXPHOS disorders it would represent a third of all cases (given that the prevalence of respiratory chain defects is calculated at 1 in 5000 live births). A caveat is that these conclusions assume c.815-27T>C is pathogenic in its own right. However, the low levels of wild-type *NUBPL* transcript may allow for sufficient complex I levels for normal function. Such high frequencies could occur due to genetic drift or there may be some selective advantage conveyed by the mutation, for instance related to adaptation to the cold climate. Taken together these data support the idea that mutations in *NUBPL* are likely deleterious.





**Figure 3.3: Expression of variant Ind1 proteins in *Yarrowia lipolytica***

**(A)** SDS-PAGE followed by immunolabelling of Ind1 in protein samples of mitochondrial membranes from strains expressing wild-type Ind1 and the indicated protein variants, in the *ind1Δ* background. Antibodies against the mitochondrial protein aconitase (Aco1) and subunit 2 of succinate dehydrogenase (Sdh2) were used to confirm equal loading and transfer.

**(B)** Transcript levels of the L102P variant compared to wild-type *IND1* by RT-PCR. A PCR reaction for *PYC1/ YAL10\_C24101g*, containing an intron, showed that the cDNA samples were free from genomic DNA and equal in cDNA content. *ACT1* was used as an additional control for the amount of cDNA template.

### 3.2.3 Protein stability of Ind1 variants

In order to study the effect of protein variants on *NUBPL*, the equivalent change was made on the corresponding residue in *Yarrowia lipolytica* Ind1. The five mutations outlined above were introduced by site-directed mutagenesis into a plasmid carrying a strep-tagged version of the *IND1* gene under the control of its native promoter and transformed into *Yarrowia lipolytica* strain GB10 with a genomic deletion of the *IND1* gene (Bych et al. 2008). This strain was previously engineered to contain the *NDH2i* gene, encoding an alternative NADH dehydrogenase targeted to the matrix side of the inner mitochondrial membrane which allows the strain to bypass the absolute requirement of respiratory complex I (Kerscher et al. 2001). As mentioned, the previously reported branch-site mutation resulting in human D273QfsX31, and *Yarrowia lipolytica* N271QfsX31, was included for comparison (Wydro & Balk 2013).

In order to test whether any of the protein variants affect protein stability, mitochondrial membranes of each strain were isolated and subjected to Western blot analysis to visualize Ind1 protein levels. Protein expression from a gene in a plasmid, even under its native promoter, will likely differ from a stable chromosomal expression, for example expression levels could be many times higher. This could create an artificial context in which to compare protein stability, where the levels of expressed proteins exceed the constitutive machinery needed for protein degradation. Therefore, care should be taken not to over-interpret the results. Here, western blot analysis can only show whether or not the proteins have been completely degraded, and act as an additional control that the Ind1 variants are properly expressed, but comparing relative levels may not be relevant to the native context. The D103Y, G136D, L191F and G285C variants displayed levels of Ind1 protein similar to *ind1Δ* + *IND1* (henceforth referred to as complemented wild type, cWT) (Figure 3.3A). In humans, analysis of patient fibroblasts with variant D105Y showed a substantial decrease in *NUBPL* protein levels (Kevelam et al. 2013). The L102P substitution resulted in lower levels of Ind1 protein (Figure 3.3A), to approximately 45% of cWT levels. The decrease in L102P protein was similar to that seen in N271QfsX31. The N271QfsX31 protein is of lower molecular weight than Ind1 protein due to the 11 amino acid-truncation of the C-terminus. Antibodies against Aco1 and Sdh2 were used as loading controls. Additionally, the Sdh2 levels confirm that complex II subunit protein levels are unaffected in Ind1 mutants as previously reported (Bych et al. 2008).

In order to determine the reason for low levels of the L102P protein variant, RT-PCR was carried out to assess the expression of *IND1* (Figure 3.3B). Normal levels of *IND1*

transcript were observed in the L102P variant, indicating that the lower L102P protein levels are due to post translational degradation rather than decreased transcript levels. *ACT1* transcript levels were used as a control for the amount of cDNA template. A PCR reaction for *PYC1/YAL10\_C24101g* which contains an intron, was performed to show that the cDNA was free from genomic DNA and equal in cDNA content. Overall these data suggest that, with the exception of L102P, the Ind1 variants do not affect protein stability.

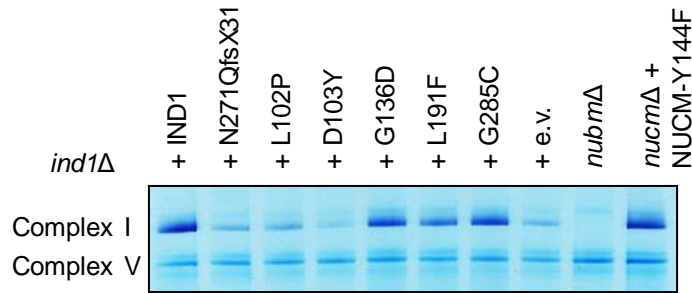
### **3.2.4 Variants L102P, D103Y and L191F are unable to support cWT levels of complex I assembly**

It has previously been shown that *IND1* knockout strains have approximately 28% of intact, active, complex I compared to wild type (Bych et al. 2008). To investigate the effect of the amino acid variants in Ind1 on complex I assembly, mitochondrial membranes were isolated and complex I content and activity assayed.

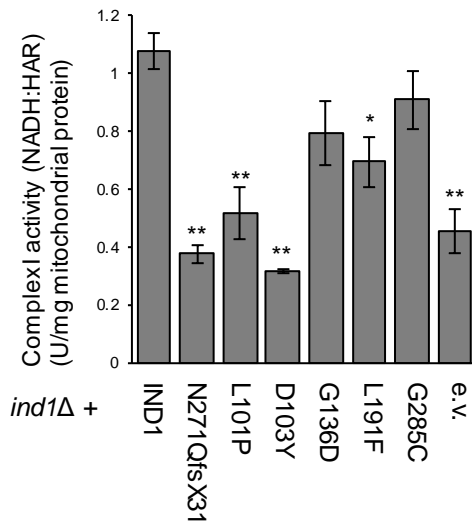
To visualise complex I content mitochondrial membrane respiratory complexes were separated using blue-native PAGE (BN-PAGE) and stained for dehydrogenase activities with NADH and nitro-blue tetrazolium (NBT). This histochemical enzyme activity stain is specific to complex I and allows easy visualisation of complex I levels. It is used as a proxy to estimate complex I levels, but not complex I enzyme activities (Sabar et al. 2005). Complex I levels were strongly decreased in L102P, D102Y and N271QfsX31, similar to the levels in the *ind1Δ* empty vector control (e.v.) (Figure 3.4A). L191F had a slight decrease in complex I levels, whereas G138D and G285C were indistinguishable from cWT. For comparison, a knockout strain of the NUBM (*nubmΔ*) subunit of complex I displayed no complex I. Another complex I mutant, where a key catalytic residue of the NUCM subunit has been mutated (Y144F) displayed complex I levels similar to cWT (Tocilescu et al. 2010). This is consistent with a mutation that affects catalysis, but does not inhibit complex I assembly. Complex V levels were unaffected by mutations in *IND1*, as has been previously reported in the *ind1Δ* mutant (Figure 3.4A).

To measure oxidoreductase activity of the peripheral arm of complex I, two different spectrophotometric assays were used. Electron transfer from NADH to the artificial electron acceptor hexaammineruthenium(III) chloride (HAR) involves only the primary electron acceptor FMN bound to the NUBM subunit of complex I and so serves as a proxy for complex I content. The L102P, D103Y, L191F, and N271QfsX31 variants showed a significant decrease in NADH:HAR activity, and therefore in complex I levels (Figure 3.4B, Table 3.5). The decrease seen in L102P, D103Y and N271QfsX31 is

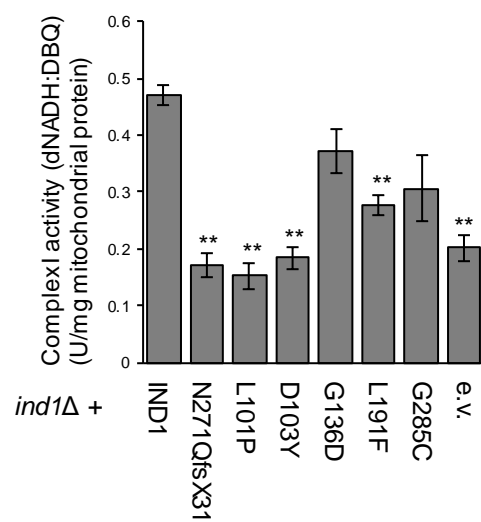
**A**



**B**



**C**



**Figure 3.4: Complex I levels and oxidoreductase activity in Ind1 variants**

**(A)** Complex I levels shown by BN-PAGE analysis of mitochondrial membrane samples from *Yarrowia lipolytica* *ind1Δ* expressing *IND1* and the indicated protein variants, compared to two well characterized complex I mutants (*nubmΔ* and *nucmΔ* + NUCM-Y144F). Complex I activity was visualised with NADH/NBT staining.

**(B)** NADH:HAR and **(C)** dNADH:DBQ oxidoreductase activity in mitochondrial membrane samples from the indicated strains. n=3 biological replicates. Error bars represent s.e.m. Two-sample t-test: \* p<0.05, \*\* p<0.01.

**Table 3.5: Complex I activity assays for Ind1 variants**

Activities are given as mean  $\pm$  s.e.m (n=3). Two-sample t-test

Protein variant	NADH:HAR activity			dNADH:DBQ activity		
	Umg <sup>-1</sup>	p-value	% of WT	Umg <sup>-1</sup>	p-value	% of WT
Ind1 (wild type)	1.07 $\pm$ 0.06	-	100	0.47 $\pm$ 0.02	-	100
e.v.	0.45 $\pm$ 0.08	p= 0.003 (**)	42	0.20 $\pm$ 0.02	p< 0.001 (**)	43
N271QfsX31	0.37 $\pm$ 0.03	p< 0.001 (**)	34	0.17 $\pm$ 0.02	p< 0.001 (**)	36
L102P	0.51 $\pm$ 0.09	p= 0.007 (**)	48	0.15 $\pm$ 0.02	p< 0.001 (**)	32
D103Y	0.31 $\pm$ 0.01	p= 0.006 (**)	29	0.18 $\pm$ 0.02	p< 0.001 (**)	39
G136D	0.79 $\pm$ 0.11	p= 0.088	74	0.37 $\pm$ 0.04	p= 0.084	79
L191F	0.69 $\pm$ 0.09	p= 0.024 (*)	65	0.28 $\pm$ 0.02	p= 0.002 (**)	59
G285C	0.91 $\pm$ 0.10	p= 0.222	84	0.31 $\pm$ 0.06	p= 0.056	65
M180V	1.02 $\pm$ 0.13	p= 0.705	95	-	-	-
M180A	1.02 $\pm$ 0.07	p= 0.780	95	-	-	-
<i>nucm</i> $\Delta$ + Y144F	-	-	-	0.04 $\pm$ 0.01	p< 0.001 (**)	9

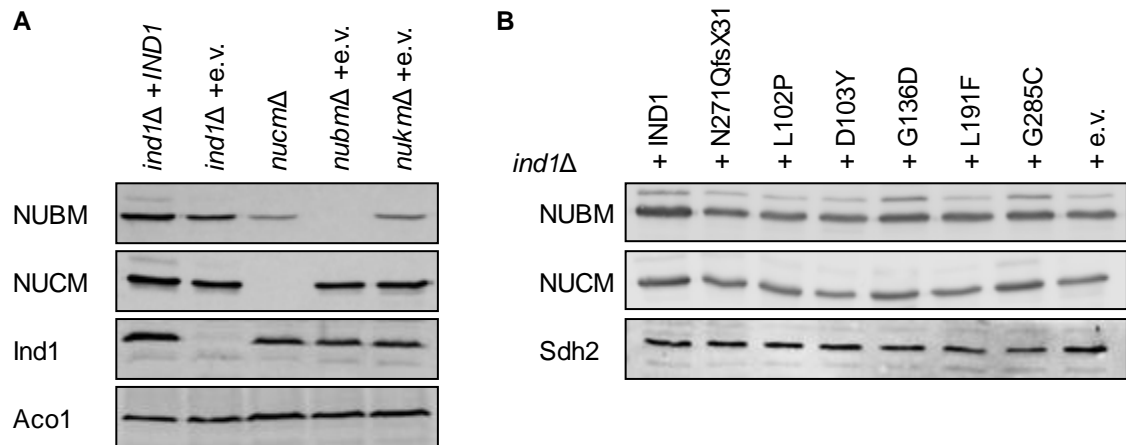
similar to the decrease in the empty-vector control. G136D and G285C are not significantly decreased compared to cWT. These data are consistent with what is found by NADH/NBT activity staining of BN-polyacrylamide gels.

The second assay involves electron transfer from deamino-NADH (dNADH) to *n*-decylubiquinone (DBQ) which encompasses all cofactors in the matrix arm of complex I and is therefore a more accurate measure of the physiological NADH:ubiquinone oxidoreductase activity. A significant decrease in electron transfer was seen with the L102P, D103Y, L191F, and N271QfsX31 variants, but not in G136D and G285C (Figure 3.4C, Table 3.5). The previously studied complex I mutant *nucm* $\Delta$  + NUCM-Y144F, which is defective in a key catalytic residue in the Q pocket, had low dNADH:DBQ activity (9% of cWT, Table 3.5, 5% of WT Tocilescu et al. 2010), despite wild-type levels of fully assembled complex I (Figure 3.4A). This is consistent with a mutant that is affected in catalysis, but not complex I assembly. Comparison of dNADH:DBQ activity between the Ind1 variants and *nucm* $\Delta$  + NUCM-Y144F (Tocilescu et al. 2010) shows that in the Ind1 variants there is still substantial activity remaining. For the Ind1 variants the NADH:HAR and dNADH:DBQ activities match up, suggesting that all the complex I present is enzymatically active.

Taken together the data from three different techniques are in agreement and indicate that the L102P and D103Y mutations are severe and cause a defect in complex I assembly of similar severity as the knock-out mutant, whereas L191F only causes a mild decrease. G138D and G285C have no significant effect on complex I levels.

### **3.2.5 The levels of NUBM and NUCM, two subunits of Complex I, are not significantly altered**

In mutants of complex I assembly, regulatory steps may be taken to either down-regulate subunit transcription and translation or to post-translationally degrade already synthesised subunits. For example, in human cells RNAi knockdown of *NUBPL* resulted in a major decrease in *NDUFV1* (Human homolog of NUBM) (Sheftel et al. 2009). Western blot analysis was carried out on mitochondrial membranes to probe the presence of complex I subunits. However, I was limited by available antibodies to subunits NUBM and NUCM. In *Yarrowia lipolytica*, there is a slight decrease in NUBM levels in the *ind1* $\Delta$  mutant compared to cWT. However, in *nucm* $\Delta$  and *nukm* $\Delta$  mutants, both subunits from the Q module, there is a greater decrease in the levels of NUBM. These data suggest that if the Q-module is not assembled, NUBM is degraded, assuming its transcription and mitochondrial import is unaffected. In contrast, the levels of NUCM



**Figure 3.5: Complex I subunits in complex I mutants and Ind1 variants**

**(A)** Protein levels of complex I subunits, NUBM and NUCM, and Ind1 visualised by immunolabelling of mitochondrial membrane samples separated by SDS-PAGE from *Yarrowia lipolytica* mutants as indicated. Antibodies against mitochondrial aconitase, Aco1, which was used to confirm equal loading and transfer.

**(B)** Immunolabelling of mitochondrial membrane samples from *Yarrowia lipolytica* *ind1Δ* expressing Ind1 and the indicated protein variants with antibodies against complex I subunits NUCM, NUBM, and Sdh2, which was used to confirm equal loading and transfer.

are similar in *nubm* $\Delta$ , *nukm* $\Delta$  and *ind1* $\Delta$  mutants. Therefore, NUCM is translated and imported in mitochondria but not assembled.

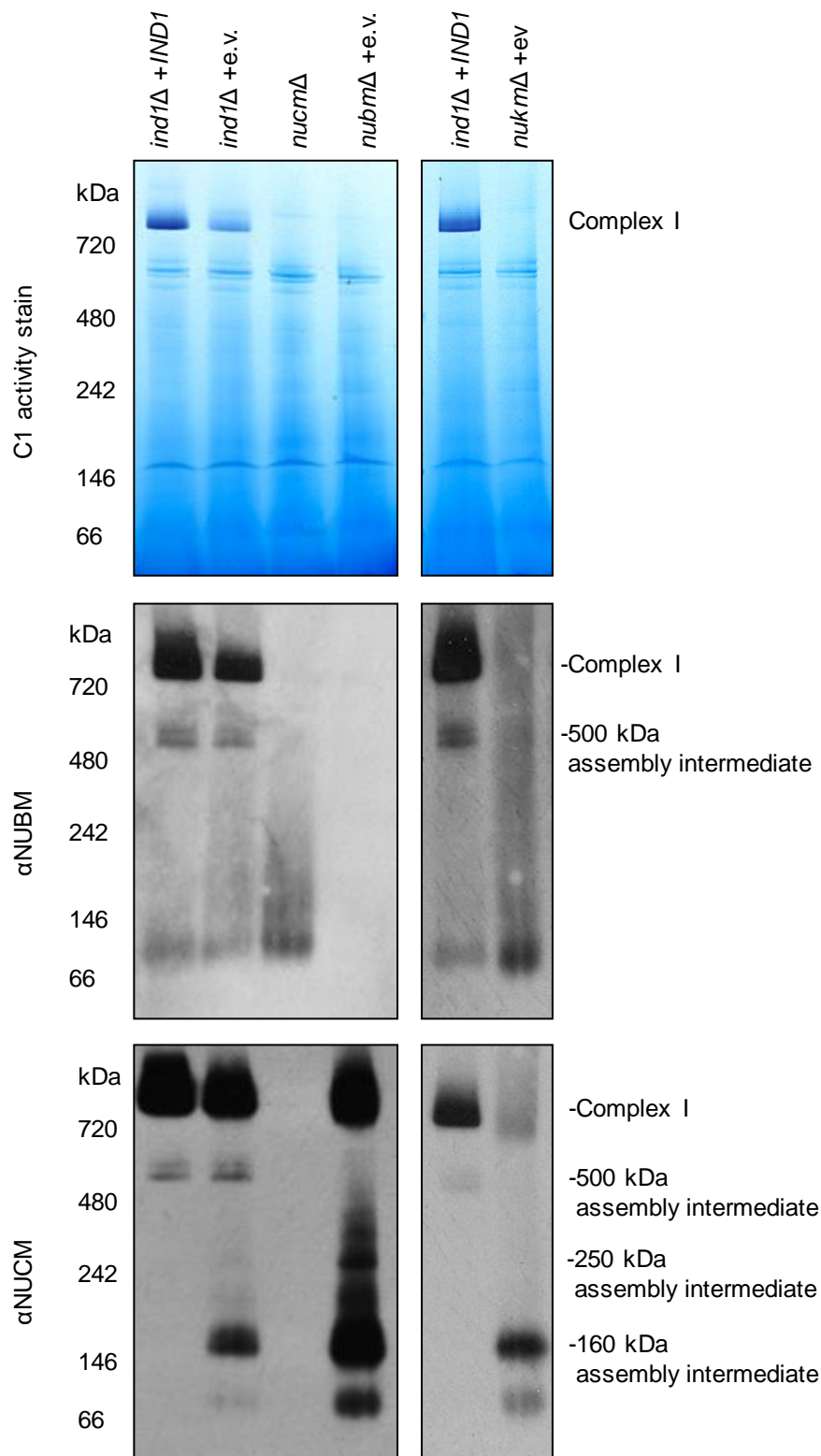
Despite an approximately 70% decrease in complex I levels in some *Ind1* variants and *ind1* $\Delta$ , the protein levels of NUCM and NUBM are largely unaffected compared to cWT, with only small decreases observed (Figure 3.5A and 3.5B). *Ind1* protein levels are slightly decreased in *nucm* $\Delta$ , *nubm* $\Delta$  and *nukm* $\Delta$  compared to cWT. These data show that the *Ind1* knock-out mutant has a minor effect on the NUBM subunit level, whereas in Q-module mutants there is a substantial decrease in NUBM protein levels. This implies *Ind1* does not affect the N module as much as the Q module does.

### 3.2.6 *ind1* $\Delta$ accumulates the Q module assembly intermediate

Given *Ind1*'s similarity to FeS scaffold proteins NBP35 and CFD1 (Bych et al. 2008) it has a suggested function inserting FeS clusters into the matrix arm of complex I. However, the exact stage at which *Ind1* acts has never been shown and the pattern of assembly intermediates has never been characterised in *Yarrowia lipolytica*. Some work has been done on other species. In *INDH* mutants, the plant homolog of *NUBPL*, assembly modules of the membrane arm of complex I are found (Wydro et al. 2013). Similarly, in humans a membrane arm assembly intermediate was observed in RNAi knockdown lines (Sheftel et al. 2009). To investigate how *Ind1* functions in complex I assembly, and to see if it is similar to that seen in plants and humans, assembly intermediate profiles were analysed in *ind1* $\Delta$  and compared to deletion strains of three well characterised complex I subunits: the N-module subunit NUBM and the Q-module subunits NUCM and NUKM.

Mitochondrial membranes were isolated from cWT, *ind1* $\Delta$ , *nubm* $\Delta$ , *nucm* $\Delta$  and *nukm* $\Delta$  and the respiratory complexes were separated using BN-PAGE. Gels were then stained for NADH/NBT activity to visualise complex I levels, or Western blotting and immunolabelling carried out with antibodies against NUBM and NUCM. In *ind1* $\Delta$ , an assembly intermediate was seen at approximately 500 kDa which was also seen in cWT (Figure 3.6). The levels of this intermediate appear equal between *ind1* $\Delta$  and cWT in figure 3.6, although in some repeats the levels are higher in cWT. This intermediate contains the NUBM and NUCM subunits and is therefore likely to contain the full matrix arm of complex I (Vogel et al. 2007), perhaps as well as the partial membrane arm. An additional assembly intermediate containing NUCM, but not NUBM, was seen at approximately 160 kDa. This intermediate is not found in cWT and likely corresponds to the Q module. A similar Q module assembly intermediate has been previously observed





**Figure 3.6: Complex I assembly intermediates in *Ind1* and Complex I mutants**

BN-PAGE analysis of mitochondrial membrane samples from *Yarrowia lipolytica* *ind1Δ* and complex I mutants; *nucmΔ*, *nubmΔ* and *nukmΔ*. Top panels: Complex I activity was visualised with NADH/NBT staining. Middle and bottom panels: Western blotting and immunolabelling with antibodies against NUCM and NUBM.

in *Yarrowia lipolytica* (Kmita et al. 2015) and contains the NUCM (human NDUFS2, bovine 49 kDa), NUGM (NDUFS3/30 kDa), NUKM (NDUFS7/PSST) and NUFM (NDUFA5/ B13) subunit. Combined these subunits would have a molecular weight of 116.42 kDa (calculated based on molecular weights in Angerer et al. 2011). Assembly factors would likely also be bound to this structure, resulting in a higher observed molecular weight ~160 kDa. Additionally, signal was seen around 80 kDa for both antibodies but this most likely represents monomer and dimers and initial assembly products. These data suggest that while *ind1Δ* can assemble some fully assembled complex I and a 500 kDa intermediate, there is a large accumulation of the Q-module intermediate.

Signal can be seen at the molecular weight corresponding to mature complex I when probing with the NUCM antibody in the *nubmΔ* strain. This is unexpected given that no mature complex I can be seen in this strain using NADH/NBT staining. However, long exposures are needed to detect the assembly intermediates. Shorter exposures show that this signal is less strong than that in cWT and *ind1Δ* as well as being at a slightly lower molecular weight. It is possible that a small amount of mature complex I is formed, below the detection limit of NADH/NBT staining.

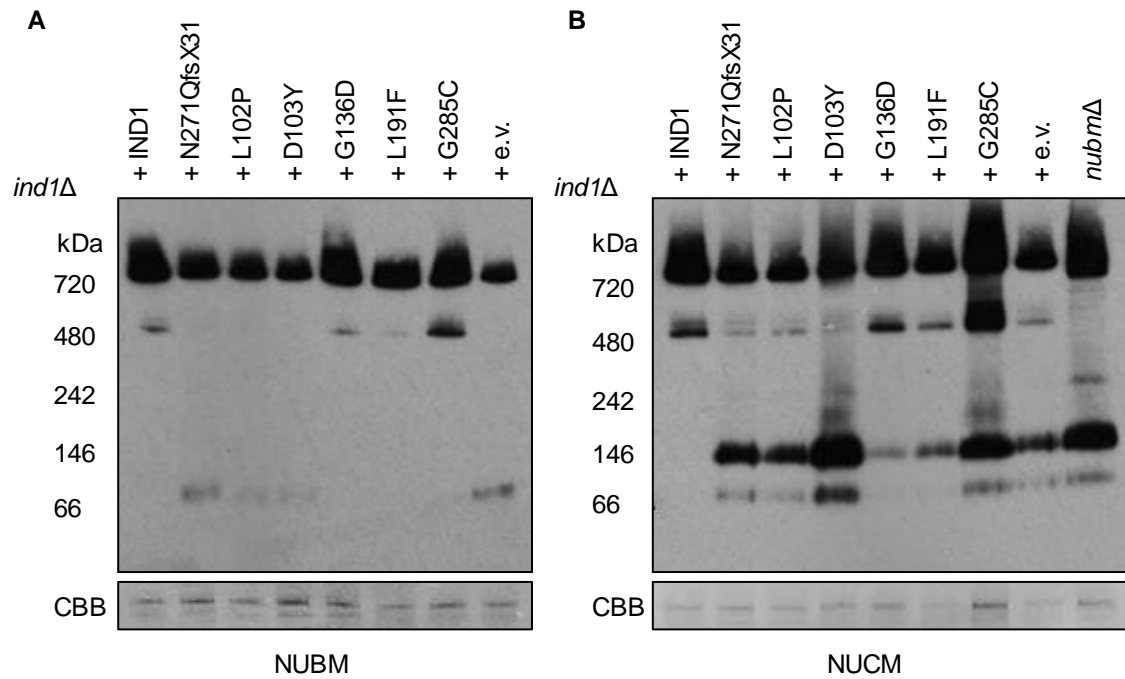
The Q module intermediate can also be seen in *nubmΔ* and *nukmΔ*, which did not contain the 500 kDa assembly intermediate (Figure 3.6). This is consistent because the Q module cannot dock to N-module if NUBM is lacking, or to the membrane module if NUKM is lacking. Therefore, a Q module, but no other, intermediate is formed.

These data show there is a clear difference between the assembly intermediates seen in *ind1Δ* and those seen in *nubmΔ*, *nucmΔ* and *nukmΔ*.

### **3.2.7 All Ind1 variants accumulate the Q module of complex I**

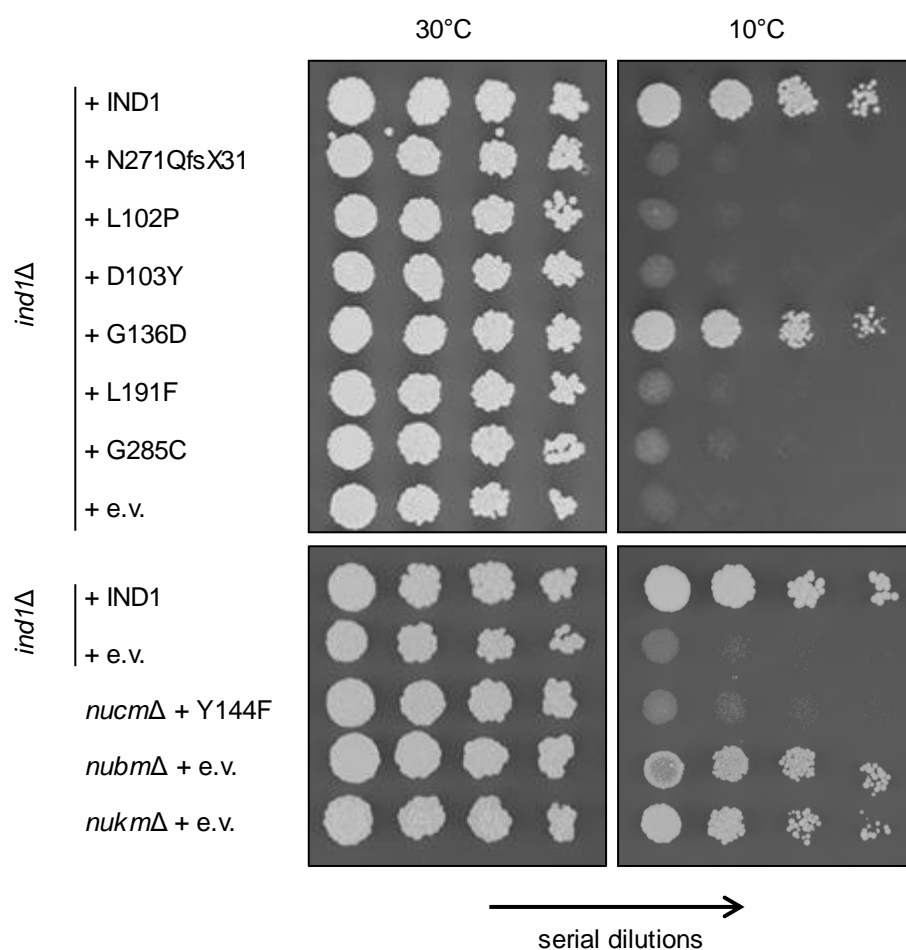
Having investigated the assembly defects in *ind1Δ*, I wanted to investigate the effect of the amino acid variants in Ind1 on complex I assembly. Complex I assembly profiles were analysed in Ind1 variants using BN-PAGE and Western blotting using antibodies against NUBM and NUCM.

The variants, even those with normal complex I levels, have the same pattern of assembly intermediates containing NUCM as the empty vector control, namely accumulation of the Q-module containing NUCM, with some minor variations (Figure 3.7A). The G136D and L191F variants accumulated low levels of the Q-module. D103Y accumulated high levels of Q module and no, or little, 500 kDa intermediate. For



**Figure 3.7: Complex I assembly intermediates in Ind1 variants**

BN-PAGE analysis followed by Western blotting and immunolabelling with antibodies against **(A)** NUBM and **(B)** NUCM of mitochondrial membrane samples from *Yarrowia lipolytica* *ind1Δ* expressing *IND1* and indicated Ind1 variants. Coomassie staining of complex V was used as a loading control.



**Figure 3.8: *Yarrowia lipolytica* growth at low temperature**

Growth of *Yarrowia lipolytica* *ind1Δ* expressing *IND1* and the indicated protein variants, and complex I mutants, *nucmΔ* + NUCM-Y144F, *nubmΔ* + e.v. and *nukmΔ* + e.v. at normal and cold temperatures. Images show a serial dilution of cultures that were spotted onto agar plates.

intermediates containing the NUBM subunit, variants with low complex I levels (N271QfsX31, L102P and D103Y) closely resembled the empty vector control, whereas those with near normal complex I levels (G136D, L191F, G285C) more closely resemble cWT (Figure 3.7B). These data show that all Ind1 variants display assembly defects, as they all accumulate Q module intermediates.

### 3.2.8 Ind1 variants are sensitive to cold

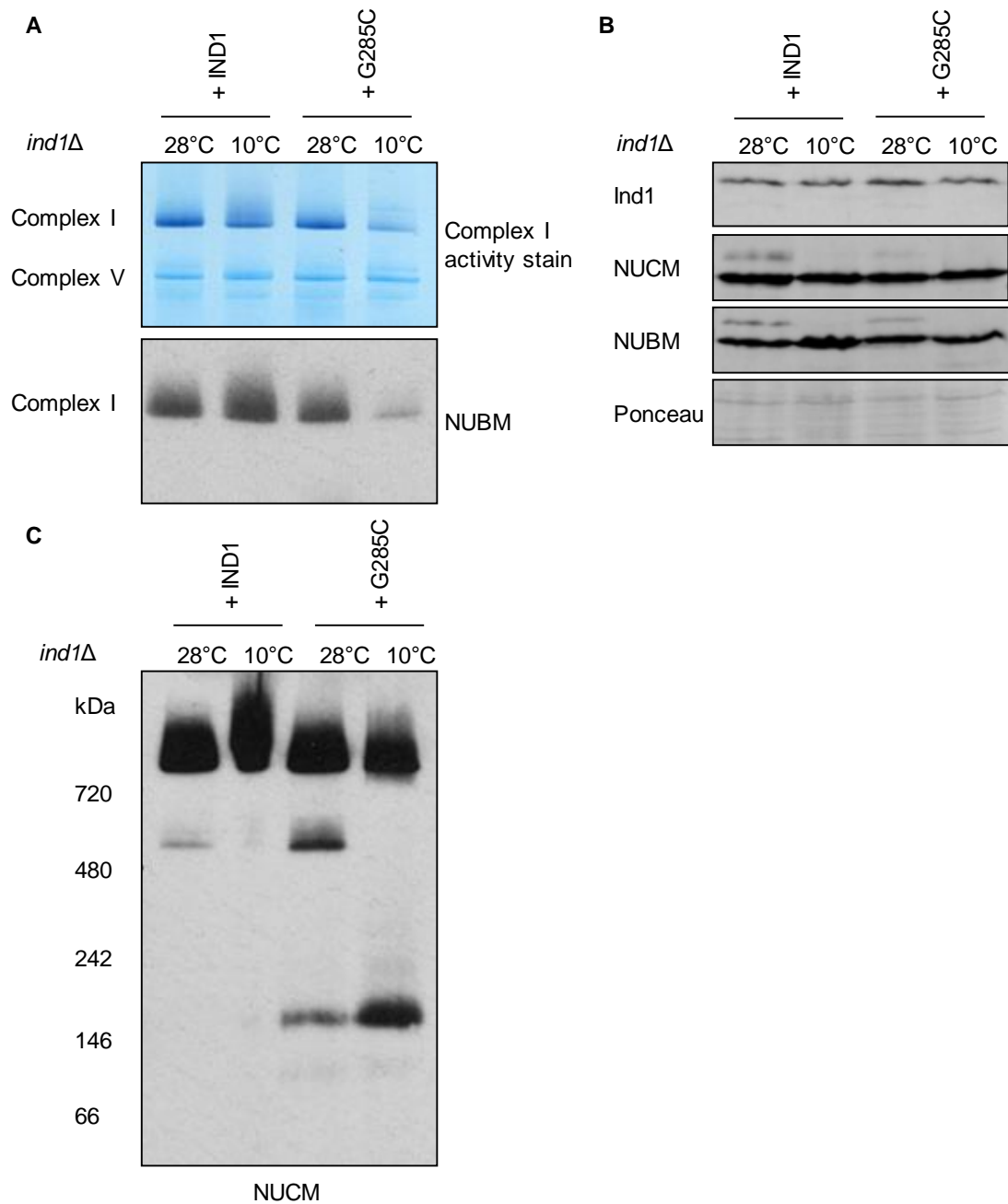
During storage of *Yarrowia lipolytica* strains on agar plates in the cold-room, it was noticed that the *ind1Δ* mutant was sensitive to cold, exacerbating the small growth difference between it and the wild-type strain. We therefore tested the effect of cold (10°C) on the growth of Ind1 variants and different complex I mutants.

*Yarrowia lipolytica ind1Δ* expressing WT Ind1 or Ind1 variants, as well as *nucmΔ* + NUCM-Y144F, *nubmΔ* and *nukmΔ*, were grown in liquid culture before being spotted onto agar plates in a dilution series. After incubation of the plates overnight at 30°C plates were either left at 30°C for a further two days or moved to 10°C for a further 10-14 days. At 30°C all strains grew similarly, despite slight variations in colony size (Figure 3.8, left panels). However, when grown at 10°C some strains displayed dramatic growth differences (Figure 3.8, right panels). The *ind1Δ* strain expressing WT Ind1, Ind1 variant G136D, *nubmΔ* and *nukmΔ* all displayed significant growth, although slower compared to those plates grown at 30°C. However, all other strains at 10°C displayed very slow, or no, further growth. Interestingly the complex I mutants lacking fully assembled complex I, *nubmΔ* and *nukmΔ*, were able to grow in cold conditions but not *nucmΔ* + NUCM-Y144F, which has fully assembled complex I but is not enzymatically active.

These data reveal a novel growth difference of certain complex I mutants grown at normal (30°C) or cold (10°C) conditions. Ind1 variants, except for G136D, also display this cold dependent growth difference.

### 3.2.9 Complex I is unstable at cold temperatures for Ind1 variant G285C

The Ind1 variant G285C has normal levels and activity of complex I (Figure 3.4A). However, it displayed a complex I assembly pattern that is different to that of cWT (Figure 3.7B) and growth at 10°C is impaired (Figure 3.8). These observations led me to investigate the assembly of complex I in the G285C variant at cold compared to normal growth temperatures. Mitochondrial membranes were isolated from cWT and G285C



**Figure 3.9: Cold temperatures lead to complex I instability in G285C variant**

**(A)** Complex I levels in mitochondrial membrane samples from *Yarrowia lipolytica ind1Δ* strains expressing wild-type *IND1* or Ind1 variant G285C isolated after growth at 28°C or 10°C. Complex I activity was visualised by BN-PAGE followed by NADH/NBT staining (upper panel) and by Western blot and immunolabelling with antibodies against NUBM (lower panel).

**(B)** Complex I assembly intermediates from *Yarrowia lipolytica ind1Δ* strains expressing wild-type *IND1* or Ind1 variant G285C isolated after growth at 28°C or 10°C. Intermediates were visualised by BN-PAGE followed by Western blot and immunolabelling with antibodies against NUCM.

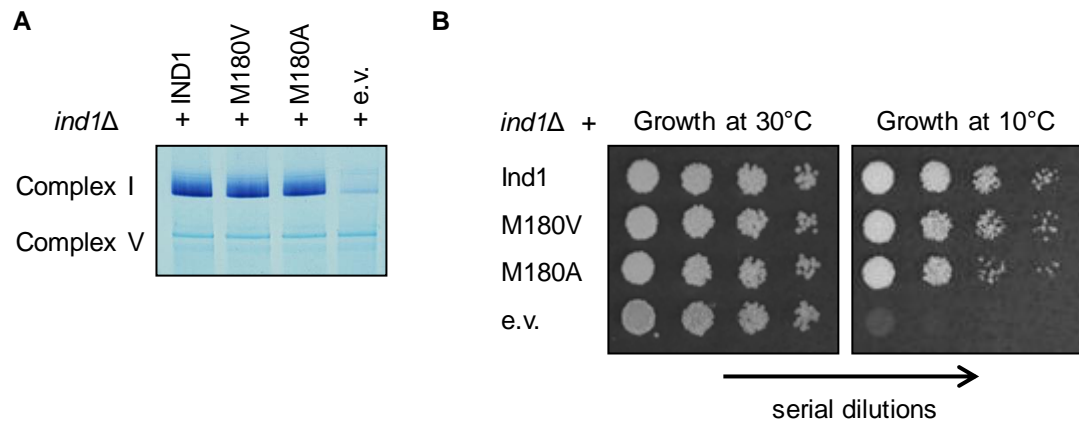
**(C)** Protein levels of Ind1 and complex I subunits in samples used in part (A) and (B). SDS-PAGE was followed by Western blot and immunostaining with antibodies against Ind1, NUCM and NUBM. Ponceau stain used to confirm equal loading and transfer.

after overnight growth at 10°C, and compared to those isolated after growth at 28°C. Respiratory complexes of the mitochondrial membranes were separated by BN-PAGE and activity stained to assess formation of complex I. cWT displayed full complex I formation in cultures grown at both 30°C and 10°C (Figure 3.9A). Conversely G285C could not support complex I assembly when grown at 10°C, with complex I levels decreased to around knock out mutant levels. BN-PAGE followed by Western blotting and immunolabelling with antibodies against complex I subunit NUBM confirmed the decrease in complex I levels in mitochondria isolated from G285C at 10°C (Figure 3.9A). SDS-PAGE and immunolabelling with the Ind1 antibody confirmed the presence of the Ind1 protein in the mitochondria isolated from G285C at 10°C, meaning the decrease in complex I levels is not due to a cold dependent degradation of the Ind1 protein (Figure 3.9B). Next, the effect of growth in cold conditions on complex I assembly intermediates was assessed. BN-PAGE followed by Western blotting and immunolabelling with antibodies against the NUCM subunit of complex I displayed different patterns between mitochondria isolated at normal and cold temperatures (Figure 3.9C). G285C samples from cold conditions are lacking a band corresponding to the assembly intermediate at 500 kDa, made up of the N and Q modules, and have an increase in the Q module intermediate at 160 kDa. This suggests at cold temperatures G285C accumulates greater levels of the Q module as it is unable to assemble it into the 500 kDa intermediate, and has lower levels of fully assembled complex I. There is still a strong signal at the size corresponding to complex I, however this is due to a long exposure time needed to see assembly intermediates.

Taken together these data show that the G285C mutant is more severely impaired in complex I assembly when grown in cold conditions. This could explain why G285C has a growth defect at 10°C similar to that of the knock out mutant.

### **3.2.10 Semi-conserved NUBPL variant V182A does not affect complex I assembly**

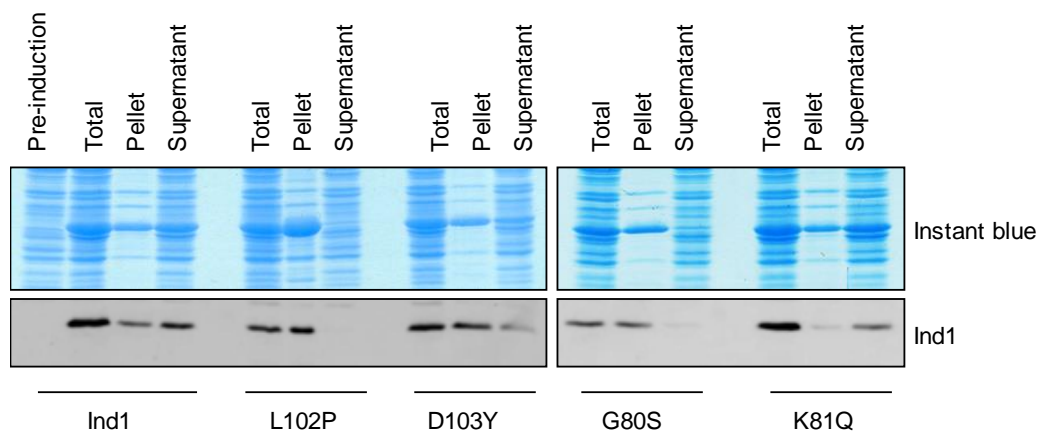
All mutations in *NUBPL* studied so far resulted in an amino acid exchange at a conserved residue in Ind1. However, for NUBPL variant V182A (Table 3.1, patient 13), there is no valine in Ind1 at position 180. Instead there is a methionine. However, these positions can be considered semi-conserved, because both valine and methionine have large hydrophobic side-chains. Alanine has a smaller hydrophobic sidechain. If M180V in Ind1 caused no deleterious effect, but a M180A did, then it would suggest that V182A in *NUBPL* is damaging.



**Figure 3.10: Complex I levels and growth in M180 Ind1 variants**

**(A)** Complex I levels in mitochondrial membrane samples from *Yarrowia lipolytica* *ind1Δ* expressing *IND1* and the indicated protein variants. Complex I activity was visualised by BN-PAGE followed by NADH/NBT staining.

**(B)** Growth of *Yarrowia lipolytica* *ind1Δ* expressing *IND1* and the indicated protein variants at normal and cold temperatures. Images show a serial dilution of cultures that were spotted onto agar plates.



**Figure 3.11: Ind1 and Ind1 variant protein expressed in *Escherichia coli***

Ind1 and Ind1 variant protein expressed in *E. coli* (*Rosetta* + pISC + pLyS). Total protein extract (T) was sonicated and centrifuged to give insoluble pellet (P) and soluble supernatant (SN). SDS-PAGE Gels were stained with instant blue (top panels) or immunolabelled with antibodies against Ind1 (bottom panels).



In order to test this hypothesis two new Ind1 variants, M180V and M180A, were created as detailed in section 3.2.3. To visualise complex I levels, mitochondrial membrane respiratory complexes were separated using BN-PAGE and stained for NADH dehydrogenase activities with NADH and NBT. Complex I levels were unaffected in both M180V and M180A (Figure 3.10A). Additionally, spectrophotometric assays using NADH and HAR showed no significant difference from cWT (Table 3.5). Growth tests under normal and cold conditions were carried out to see if any deleterious effect of M180A could be detected. However, none was found as both M180V and M180A grew similarly to cWT under cold conditions (Figure 3.10B).

These data suggest the change is not pathogenic. This is consistent with a subsequent reassessment of patient 13, who was found to possess an additional mutation leading to a chromosomal deletion (Daryl Scott, Peggy Eis, Virginia Kimonis personal communication). In that deletion were two genes, *UQCRC2* and *CDR2*, the loss of which could explain the patient's symptoms. Therefore, the mutation leading to V182A is likely not pathogenic, but rather a benign variant.

### **3.2.11 Studies on recombinant Ind1 show L102P is insoluble**

The post-translational degradation of Ind1 variant L102P, in section 3.2.3, suggests the amino acid changes could affect Ind1 protein stability. In order to assess the effects of amino acid changes on Ind1 protein stability, recombinant Ind1-strep, along with Ind1 variants, were expressed in *E. coli* (*Rosetta* strain with pISC pLys) and analysed by SDS-PAGE and Western blot analysis. After induction, total protein extract was sonicated and centrifuged to give insoluble pellet and soluble supernatant fractions. WT Ind1 was present predominantly in the supernatant fraction, but also in the pellet, indicating that most of Ind1 produced in *E. coli* is soluble (Figure 3.11). Soluble Ind1 protein is also seen with Ind1 variant D103Y, although at lower levels. However, no Ind1 protein is seen in the supernatant fraction in L102P. All protein is in the pellet fraction. This indicates that the L102P amino acid exchange causes the protein to be insoluble, possibly through mis-folding. This would suggest a reason for post translational degradation.

Two additional mutations were investigated for solubility. G80S and K81Q are located in the Walker A motif. They have both previously been shown to be deleterious, completely disabling Ind1 function (Table 3.7, Stefan Kerscher unpublished data). K81Q, similarly to WT and D103Y has Ind1 protein in both the pellet and supernatant fraction, whereas G80S, similarly to L102P, has Ind1 protein mainly in the insoluble pellet fraction (Figure 3.11).

**Table 3.6: Complex I activity in patients**

Taken from Calvo et al. 2010; Tenisch et al. 2012; Kevelam et al. 2013

Patient	CI fibroblasts (% of control)	CI muscle (% of control)	<i>Yarrowia</i> N271QfsX31 (% of WT)	<i>Yarrowia</i> Ind1 variant (% of WT)
2	27	-	34	-
3	64	-	34	29
4	56	-	34	29
6	83	67	34	65
7	19	37	34	-
10	-	31	34	-

**Table 3.7: Complex I activity assays for Ind1 variants G80S and K81Q**

Data from Dr. Stefan Kerscher

Strain	NADH:HAR activity (% of WT)	dNADH:DBQ (% of WT)
G80S	35	32
K81Q	16	17

### 3.3 Discussion

In this chapter I recreated human *NUBPL* variants in *Yarrowia lipolytica* Ind1 in order to study their effect and assign pathogenicity. Six mutations which result in amino acid changes were selected and recreated. Given the proposed function of Ind1 in complex I assembly, and that complex I deficiency is the cause of clinical symptoms, the first indicator of pathogenicity was to look at complex I levels and activity. I then discovered that looking at growth in the cold was also a good indicator of pathogenicity. Ind1 variants L102P and D103Y showed the strongest decrease in complex I, similar to levels in N271QfsX31 and the knock-out mutant. This suggests that these mutations completely abolish the function of Ind1. This is as expected given they affect highly conserved residues in the switch I motif which is required for the NTPase activity of Ind1. These results are also similar to the Ind1 G80S and K81Q variants which affect the Walker A motif (Table 3.7, Stefan Kerscher unpublished data). D103Y has wild-type levels of protein, which must all be non-functional because complex I levels are the same as in the knock-out mutant. L102P appears to cause post-translational degradation due to protein mis-folding, as the protein is insoluble. Together these data show that mutations to the P-loop are highly disruptive of function and L102P and D103Y are almost certainly pathogenic.

The L191F and G136D variants affect residues in the MRP domain. The function of the MRP domain is not well defined and so it was hard to predict the effect of the mutations. L191F resulted in a moderate decrease in complex I levels and activity, which was clearly less than cWT, but above that of the knock-out mutant. Conversely G136D appeared to have no significant effect on complex I levels and acted like cWT in almost all tests. The one exception is in assembly, where Q module assembly intermediates accumulated. In summary L191F appears to be pathogenic. G136D appears to display assembly defects, but is not necessarily strongly deleterious. G136D was not associated with a clinical case and so, out of all the variants studied, was the one least likely to be deleterious.

The G285C variant was the hardest to classify. It did not have a significant effect on complex I levels, however, it did display an accumulation of the Q module assembly intermediate. Unlike cWT, G285C was unable to grow in the cold and was found to be more severely defective in complex I assembly at this temperature. Why the assembly defect seen in G285C is exacerbated at cold temperatures remains an open question. Consequently, G285C can be classed as potentially pathogenic but requires further study.

The V182A mutation, recreated as M180V and M180A in *Yarrowia lipolytica*, displayed no deleterious effects. This patient's symptoms were subsequently found to be caused by another mutation. The lack of deleterious effect found therefore serves as a good internal control, and lends credence to the classification of other mutations as pathogenic.

Data of complex I levels exists for 5 of the 13 patients (Calvo et al. 2010; Tenisch et al. 2012; Kevelam et al. 2013) (Table 3.6). Assays were performed on either muscle biopsies or fibroblasts or both. Sometimes the result differs in the same patient depending on tissue type looked at. For example, patient 7 has 19% complex I levels in fibroblasts but 37% from muscle biopsies. These levels can be compared to what was found in the corresponding *Ind1* variants (Table 3.6). Generally, there is a poor match between levels seen in *Yarrowia lipolytica* and human. Combined with the heterogeneity between tissue type and method it is hard to draw any meaningful conclusions.

In summary four of the six mutations (L104P, D105Y, L193F and G285C) were classed as pathogenic and one (G138D) was classed as potentially pathogenic. These results allow a degree of confidence that certain patient mutations are causative of complex I disorders. Along with previous studies (Ahlers et al. 2000; Kerschler et al. 2004; Varghese et al. 2015; Gerber et al. 2016) these results reconfirm the utility of *Yarrowia lipolytica* as a model for human mitochondrial disorders. *Neurospora crassa* has also proved useful as a model (Duarte et al. 2005; Ferreira et al. 2011). However, care must be taken not to over-interpret the results as the two organisms are separated by hundreds of millions of years of evolution and so an observation in one organisms might not be relevant for the other. In my data were two examples where an effect seen in *Yarrowia lipolytica* was not seen in humans. In human *NUBPL* RNAi HeLa lines the NDUFV1 subunit is substantially decreased (Sheftel et al. 2009), whereas in *Yarrowia lipolytica ind1Δ* protein levels of NUBM are only mildly affected. Secondly the D103Y variant appears to have stable protein levels in *Yarrowia lipolytica*, but the human patients with D105Y have almost no *NUBPL* protein (Kevelam et al. 2013). Ideally these mutations should also be tested in human cell lines or patient biopsies. For example, further work should be done on G285C in human lines to see if it results in a decrease in complex I in patient tissues. However, studies in a yeast model are quicker and more cost effective, and so can be used as a first pass screen to look for potentially pathogenic mutations. An additional advantage is that *Yarrowia lipolytica* can be studied whilst haploid, meaning the effect of potentially recessive mutations can more readily be studied.

Previously published work on the role of Ind1 in complex I assembly does not contain much mechanistic detail. It is observed that in the *ind1* $\Delta$  mutant complex I levels are decreased to approximately 28% of WT (Bych et al. 2008). However, since its discovery in 2008 much work has been done on the sequence of events in the modular assembly of complex I. We have a better understanding of the assembly intermediates and which subunits they contain. It is therefore possible to revisit the question of what role Ind1 plays in complex I assembly. This will be discussed fully in the general discussion.

Little previous work has focused on growth defects associated with complex I mutants in *Yarrowia lipolytica*. I decided to investigate growth phenotypes associated with complex I and Ind1 mutants as they could prove a good proxy of the severity of mutation. Ultimately serious defects in cellular systems should manifest themselves as serious growth defects. At normal temperatures, 30°C, *Yarrowia lipolytica* containing Ind1 variants and complex I mutants grew similarly to WT, despite a slight decrease in colony size. However, after being moved to cold temperatures, approximately 10°C, there were dramatic growth differences. WT, G136D, *nubm* $\Delta$  and *nukm* $\Delta$  were able to grow, albeit at a much slower rate than at normal temperatures. Conversely all other Ind1 variants and *nucm* $\Delta$  + Y144F displayed almost no growth. The reason for these different growth phenotypes remains unclear. An initial hypothesis was that those strains without complex I could not survive in the cold, possibly as a result of an inability to maintain the proton gradient which could be used to generate heat. However, two mutants with no assembled complex I, *nubm* $\Delta$  and *nukm* $\Delta$ , grew, whereas another (*nucm* $\Delta$  + Y144F) which has fully assembled complex I, but which is catalytically dead, did not. Further investigation is needed to uncover the cause of this growth defect.

In summary using *Yarrowia lipolytica* as a model to investigate pathogenic mutations in *NUBPL* is both clinically useful and viable as a research method. In the future *Yarrowia lipolytica* would be a good system to test out mutations in numerous complex I deficiency cases, and indeed many other more varied mitochondrial disorders. It also leads to a greater understanding of the role of *NUBPL*/Ind1 in complex I assembly, as well as the discovery of a novel temperature dependent growth phenotype. However, more work is needed to understand the process of complex I assembly in *Yarrowia lipolytica*, and the differences that exist between the human and yeast pathway, before it is possible to tell how relevant findings are to human disease.

## 4 Life without complex I in a parasitic plant

### 4.1 Introduction

Parasitism is a life history strategy successfully employed across all domains of life whereby nutrition is obtained from a host (Poulin & Morand 2000). The evolution of parasitism is often associated with strong selection pressures which can drive the parasite's adaptations in new, and often unexpected, directions. An example of this is organellar genome reduction, or loss, which is a recurring phenomenon among eukaryotic parasites. For example, the intestinal parasite *Giardia lamblia* has undergone extensive mitochondrial reductive evolution, and has lost the organellar genome altogether. In plants, reductive evolution of the chloroplast genome has been studied extensively. In some plant parasites, photosynthetic genes in the chloroplast become pseudogenised, or lost, as the plant becomes less dependent on its own photosynthetic abilities, and more dependent on the photosynthate of its host. The most striking example of this is in *Rafflesia lagascae*, which may have lost its entire chloroplast genome (Molina et al. 2014).

Parasitism has evolved numerous times in plants, with current estimates suggesting it has evolved at least 12 to 13 times in angiosperms, making up perhaps as much as 1% of angiosperm species (Westwood et al. 2010). The best known and studied examples include the Orobanchaceae family, or Broomrapes, and the Rafflesiaceae family, which includes a species with the largest recorded flower, *Rafflesia arnoldii*. Parasites can be divided into two classes: facultative and obligate. Facultative parasites can complete their life cycle without a host, making parasitism an optional part of their lifecycle. Obligate parasites need a host to complete their lifecycle, and so are highly dependent on their host species. Plant parasites can also be divided into two classes, hemiparasites and holoparasites, depending on their photosynthetic capacity. Hemiparasitic species, which include both facultative and obligate species, retain their photosynthetic capacity. They often obtain other nutrition, such as water and minerals, from their hosts. Holoparasitic species lack photosynthetic capacity and therefore obtain all sugars from their host plant. Consequently, many holoparasitic plants have lost the ability to make chlorophyll. Orobanchaceae contain species across this parasitic spectrum, including facultative and obligate, hemi- and holo-parasites (Westwood et al. 2010).

The mitochondrial genomes of angiosperms are notable for their large and variable sizes (Mower et al. 2012; Sloan et al. 2012), with a 200-fold size range (Skipington et al.

2015). Plant mitochondrial genomes have many more protein coding genes than animal ones (Adams et al. 2002), for example the gymnosperm *Cycas taitungensis* has 41, and *Arabidopsis thaliana* has 31. By contrast the human mitochondrial genome has 13. Almost all Angiosperms encode a core set of 24 mitochondrial genes, including the respiratory complex subunits. A further 17 are present in some, but not all, and are mainly made up of ribosomal protein genes. Notably, many plant mitochondrial genomes encode complex II subunits, *sdh3* and *sdh4*, in contrast to humans who encode all complex II subunits in the nuclear genome. Plants encode 9 complex I subunits in the mitochondrial genome, as opposed to 7 in humans. As well as the 7 hydrophobic membrane domain subunits, plant mitogenomes additionally encode Nad7 (bovine 49 kDa subunit) and Nad9 (30 kDa subunit).

Compared to the attention parasitic plant chloroplasts have received, the mitochondrial genomes of parasitic plants have been relatively understudied. Until recently only a couple of mitogenomes from the Rafflesiaceae family had been sequenced (Xi et al. 2013; Molina et al. 2014). The most striking thing about these mitogenomes was the large degree of horizontal gene transfer. Little evidence of reductive evolution was found. However, two studies in 2015 changed this. The authors sequenced the mitochondrial genomes of four mistletoe species and discovered massive gene loss (Skippington et al. 2015; Petersen et al. 2015).

Mistletoes are hemiparasitic stem parasites from the order Santalales (Mathiasen et al. 2008). One mistletoe genus, *Viscum*, contains approximately 100 parasitic species (Petersen et al. 2015). Its best-known member is the common European mistletoe, *Viscum album*, which is widely recognisable from European folklore. The first study, (Skippington et al. 2015), sequenced the mitogenome of a Malaysian mistletoe, *Viscum scurruloideum*. It was very small in size at 66 kb, making it the smallest recorded land plant mitogenome. In contrast, *Viscum* contain some of the largest nuclear genomes in plants, up to 201 Gb (Zonneveld 2010). Strikingly, the mitogenome of *Viscum scurruloideum* has lost all mitochondrial *nad* genes, which code for subunits of complex I. Given that the wholesale transfer of multiple *nad* genes to the nuclear genome is highly unlikely, this strongly implied that complex I had been lost. The loss of complex I is unprecedented in multicellular eukaryotes. In addition to the loss of complex I coding genes, numerous other genes had been lost. Out of 24 core genes found in virtually all angiosperm mitogenomes, 11 had been lost. In addition, out of 17 genes present in many, but not all, angiosperm mitogenomes, 11 had been lost. However, the mitogenome of *Viscum scurruloideum* retained a usual complement of other respiratory complex subunits, for example the *cox* and *atp* genes, and many ribosomal RNA genes.

The mitogenome of *Viscum scurruloideum* is therefore unusual, and led the authors to describe it as “wonderfully bizarre”.

The second study, (Petersen et al. 2015), sequenced the complete mitogenome European mistletoe, *Viscum album*. This mitochondrial genome was bigger than that found in *Viscum scurruloideum*, at 565 kb. They also reported a partial genome mitogenome for two other *Viscum* species: *Viscum crassulae* and *Viscum minimum*. This second study also found the absence of all complex I encoding genes in the mitochondria. However, they did not find the usual complement of other mitochondrial genes, claiming the genes for all other respiratory complexes and rRNA were either lost or pseudogenised. This would represent a state of even higher reductive evolution from that seen in *Viscum scurruloideum*. However, the authors of the first study dispute the findings from *Viscum album*. In a subsequent reanalysis of the *Viscum album* mitogenome 21 protein coding genes were found that include a full set of respiratory complex II, III, IV and V genes, as well as rRNA genes (Skippington et al. 2017). This would suggest that both species lack complex I, but not the other respiratory complexes.

This work has four main objectives. The first is to confirm, through biochemical and proteomic techniques, that complex I has been lost from *Viscum album*. The loss needs to be confirmed as it remains possible that the *nad* genes have been transferred to the nuclear genome. The second is to determine whether complexes II to V are present or not, thereby solving the dispute in the literature. The third is to understand how *Viscum album* has adapted to life without complex I. The fourth is to study *Arabidopsis thaliana* complex I mutants to see what lessons can be learned about coping with low levels of complex I.

## 4.2 Results

### 4.2.1 Isolation of mitochondria from *Viscum album*

Mitochondria from *Viscum album* have, to the best of my knowledge, only been isolated once before (Petersen et al. 2015). However, these were used for sequencing rather than biochemical studies, and so yield was not a consideration. Therefore, mitochondrial purification protocols needed to be developed. Existing protocols for plant mitochondrial purification were tried and optimised in an iterative fashion.

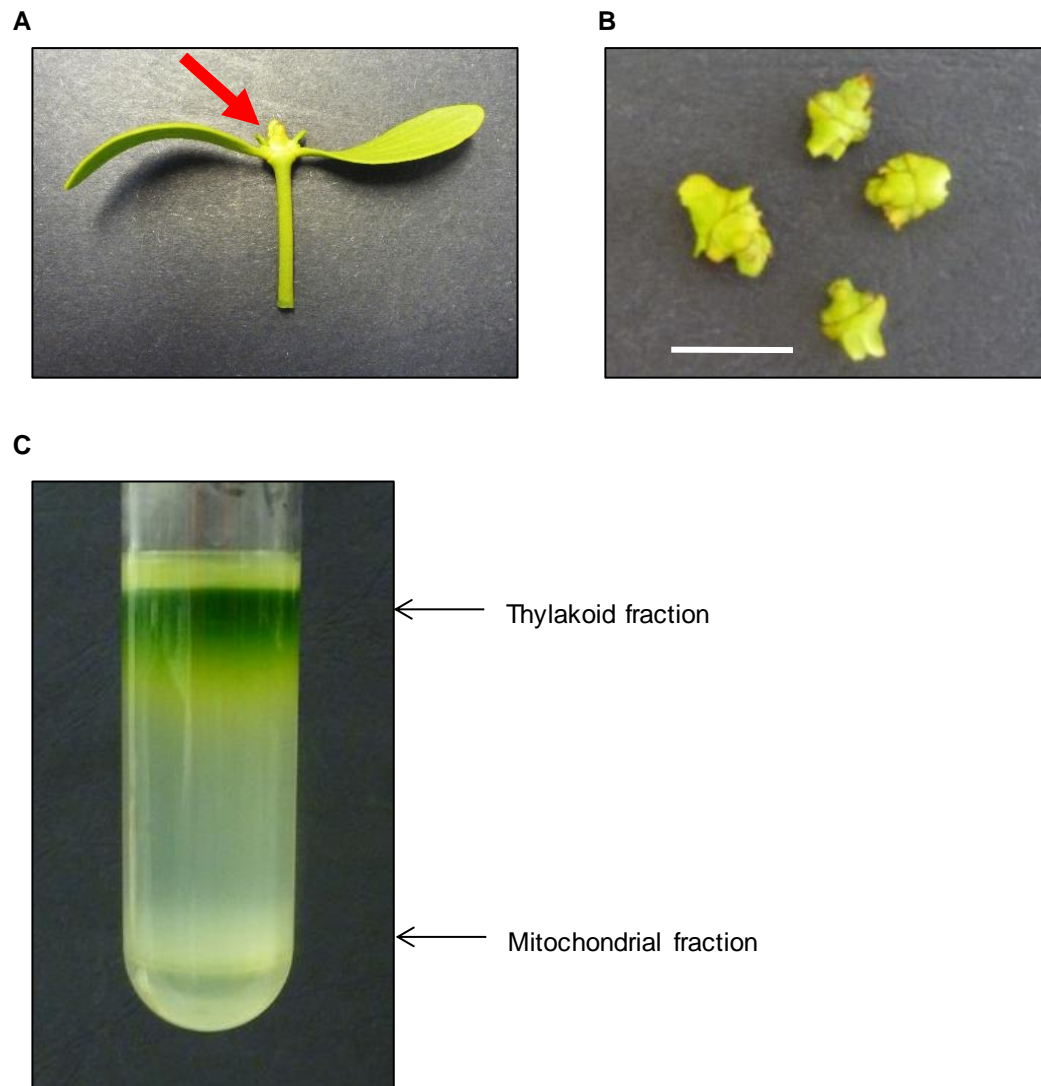
Material was collected from three separate *Viscum album* plants from three locations (Table 4.1) (Collected by Dr. Janneke Balk). First, a mitochondrial purification on leaves based on the protocol from (Sweetlove et al. 2007) was attempted. However due to the



presence of chloroplasts, and the viscous nature of the extract, no mitochondria could be extracted. Next, a purification was attempted on the leaf buds (Figure 4.1A, B). The purification was based on a modified protocol from (León et al. 2007). Briefly, buds were homogenised using a pestle and mortar and glass beads. Differential centrifugation was then performed resulting in a pellet which contained both green (chloroplast) and brown (possible mitochondria) parts. This suggested that a mitochondria-rich fraction had been isolated, albeit with chloroplastic contamination. A sample was taken here, and will be referred to from now on as “crude” mitochondria. A purer mitochondrial fraction was obtained by taking the crude fraction and separating out mitochondria and thylakoid membrane vesicles from chloroplasts on a Percoll gradient, according to (Sweetlove et al. 2007) (Figure 4.1C). These mitochondria will be referred to as “Percoll-purified” mitochondria.

In order to confirm that mitochondria had been purified, Western blot analysis with antibodies against the mitochondrial marker porin (also known as Voltage-Dependent Anion Channel (VDAC)), was performed. Porin is a good mitochondrial marker as it is highly conserved, making detection more likely. Both total protein from the bud and putative mitochondrial fraction were analysed, along with mitochondria from *Arabidopsis thaliana* callus, to act as a control. Faint bands were seen in the total protein extract but these were much stronger in the putative mitochondria fraction, both for crude and Percoll-purified mitochondria (Figure 4.2A, B). Two bands were apparent, representing different porin isoforms. These were of a higher apparent molecular weight than the main porin isoform in *Arabidopsis thaliana*. The enrichment of porin in the mitochondrial fraction suggest that mitochondria had been isolated from *Viscum album*.

Further Western blot analysis was performed to see if various mitochondrial proteins were present in *Viscum album*, for example *cox2* subunit of complex IV and the complex I assembly factor, INDH. No bands were observed in the *Viscum album* sample (Figure 4.2C, D). However, the *cox2* gene is present in the mitogenome and, as seen in section 4.2.4, *cox2* protein is found in *Viscum album*. It is therefore likely that the antibody, which is specific for *Arabidopsis thaliana*, is unable to recognise *Viscum album* *cox2*. This may be due to high levels of sequence divergence in *Viscum album*, for example *Arabidopsis thaliana* and *Viscum album* *cox2* are 67% identical at the protein level, whereas *Arabidopsis thaliana* and *Zea mays* (a more evolutionarily distant species) *cox2* are 88% identical. Additionally, no INDH protein was detected in *Viscum album*. Western blot analysis was not continued, because it was not possible to distinguish if an absent band meant a protein was missing or there was an antibody recognition issue.



**Figure 4.1: Purification of mitochondria from *Viscum album***

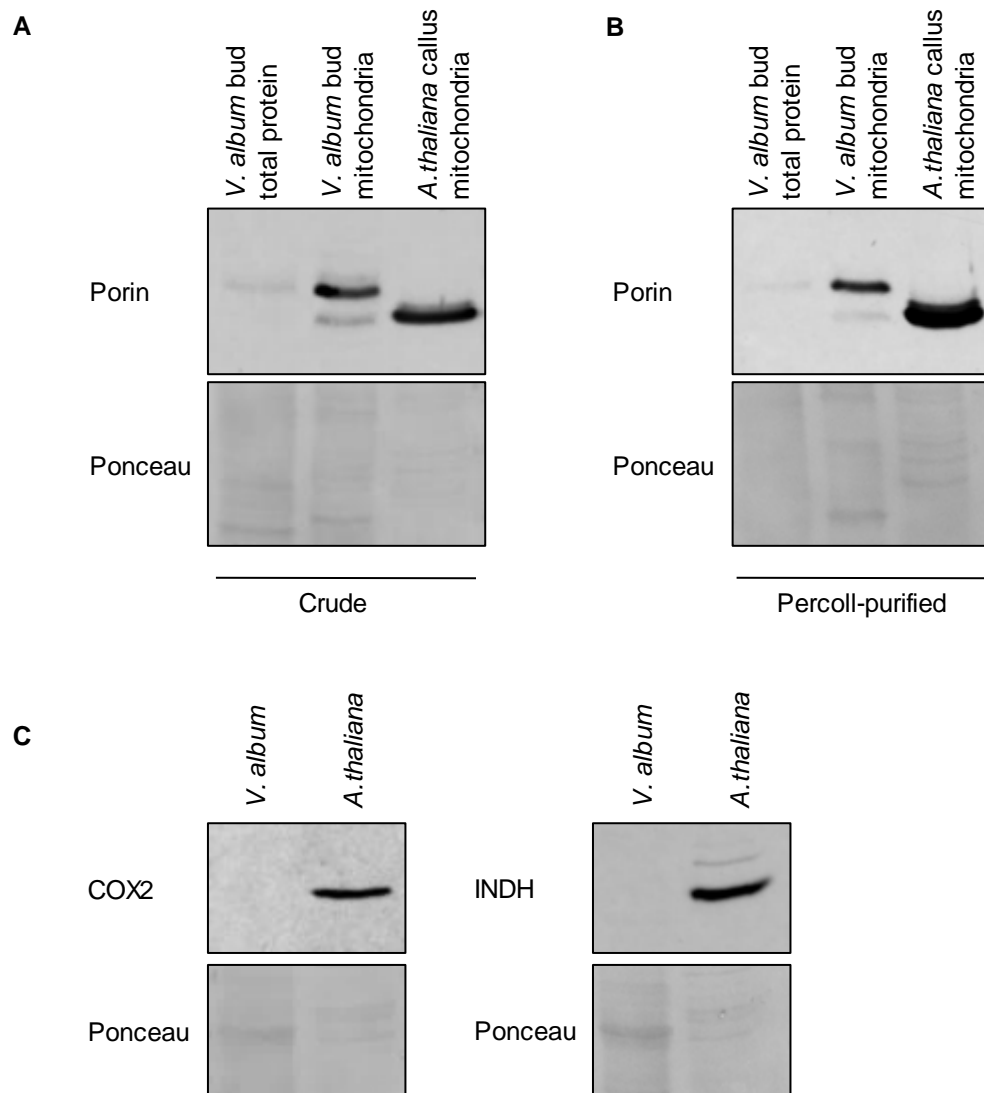
(A) *Viscum album* leaves and leaf bud (indicated with red arrow).

(B) Leaf buds used in mitochondria purification. Scale bar = 1 cm.

(C) Percoll gradient after centrifugation. *Viscum album* sample has been separated into thylakoid and mitochondrial fraction.

**Table 4.1: *Viscum album* sample collection**

Plant number	Date harvested	Elevation from sea level (m)	Elevation from ground (m)	Host plant
1	26/3/2017	15	5.5	<i>Acer campestre</i>
2	17/4/2017	10	2	<i>Acer saccharum</i>
3	17/4/2017	10	3	<i>Malus pumila</i> "Bramley"



**Figure 4.2: Western blots of *Viscum album* mitochondrial purifications**

(A) and (B) SDS-PAGE followed by immunolabelling of porin in *Viscum album* total protein and mitochondrial samples, and *Arabidopsis thaliana* callus mitochondria. Ponceau stain used to confirm equal loading and transfer. Mitochondrial samples from crude purification (A) and Percoll purification (B).

(C) SDS-PAGE followed by immunolabelling of cox2 in *Viscum album* crude mitochondria and *Arabidopsis thaliana* callus mitochondria.

(D) SDS-PAGE followed by immunolabelling of INDH in *Viscum album* crude mitochondria and *Arabidopsis thaliana* callus mitochondria.

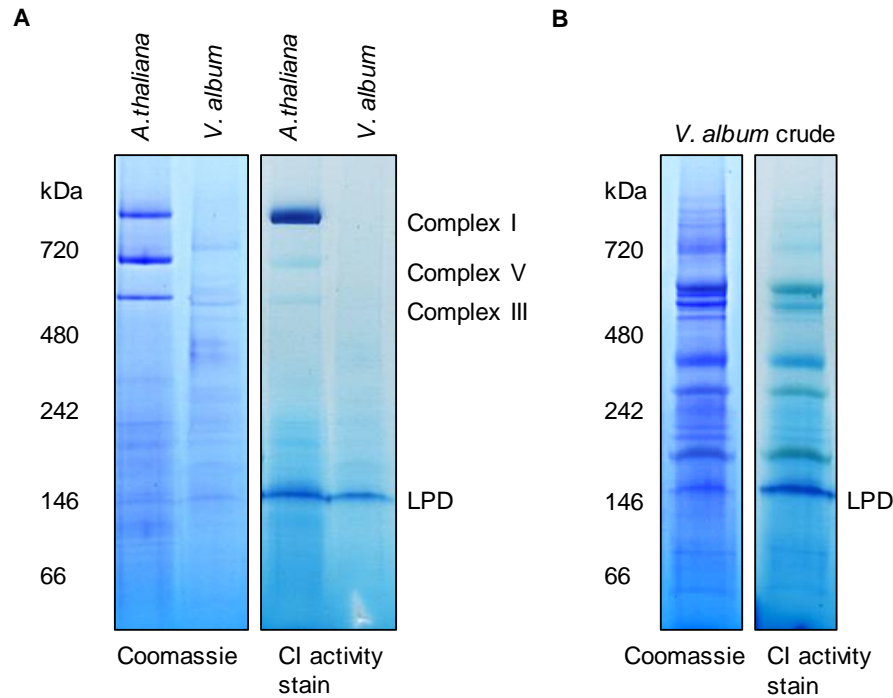
#### **4.2.2 BN-PAGE and in gel-activity assay indicate *Viscum album* lacks complex I**

Having purified mitochondria from *Viscum album*, the next stage was to determine the state of the respiratory chain. In order to visualise complex I content of *Viscum album* mitochondrial complexes were separated using BN-PAGE. They were then stained with Coomassie, to visualise protein content, or NADH and NBT, to visualise complex I levels (Sabar et al. 2005). This was compared to *Arabidopsis thaliana* callus mitochondria samples (Figure 4.3A). Callus provide a good comparison to buds, as both are meristematic tissues, with a high energy demand, and therefore contain high levels of mitochondria. In both crude and Percoll-purified *Viscum album* mitochondria an activity band was seen at around 146 kDa. This band was also seen in the *Arabidopsis thaliana* sample. This activity stain is caused by the dihydrolipoamide dehydrogenase (also known as lipoamide dehydrogenase) subunit of pyruvate dehydrogenase complex, which also has NADH dehydrogenase activity (Meyer et al. 2009) and can act as an additional confirmation that mitochondria have been isolated, as well as a loading control. These NADH/NBT activities were not seen in the Coomassie-stained sample, showing the specificity of the activity stain. In the *Arabidopsis thaliana* sample a clear activity band can be seen above 720 kDa, which represents respiratory complex I. This band is completely absent in both crude and Percoll purified mitochondria from *Viscum album* (Figure 4.3A, B). This indicates an absence of respiratory complex I in *Viscum album*. In the crude mitochondrial samples, additional green bands can be seen, which are absent from the Percoll-purified mitochondria. These are contamination from chloroplastic complexes. None of these display any activity, and so do not alter the results. These data strongly suggest that *Viscum album* does not contain any respiratory complex I.

#### **4.2.3 BN-PAGE and in gel-activity assays suggest *Viscum album* retains complex II, III and IV**

In- gel activity assays indicate that complex I is indeed absent in *Viscum album*. The next question is whether the other respiratory complexes are present, as suggested in (Skippington et al. 2015; Skippington et al. 2017), or are absent, representing complete loss of the respiratory chain as suggested in (Petersen et al. 2015). In order to assess the intactness of the respiratory chain, various BN-PAGE and in-gel activity stains were performed.

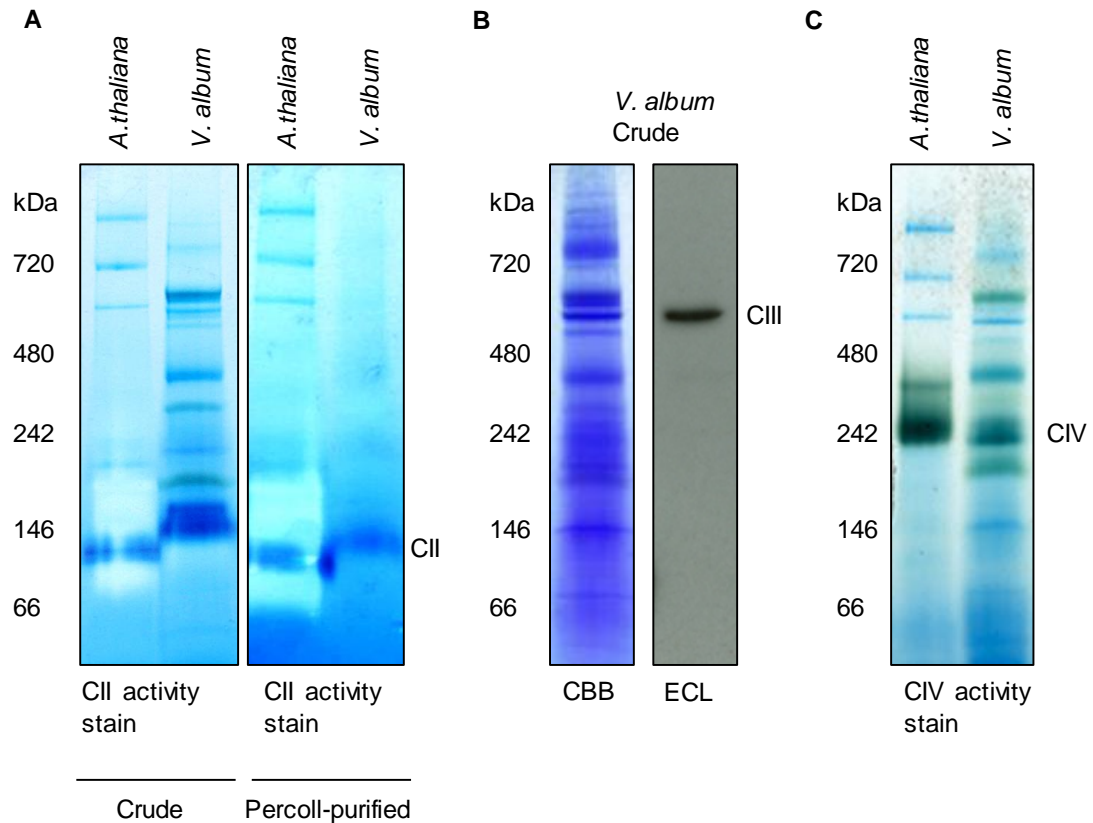
To visualise complex II, *Arabidopsis thaliana* and *Viscum album* mitochondria were separated by BN-PAGE and activity stained with PMS and NBT (Sabar et al. 2005). A



**Figure 4.3: Complex I activity stain is absent in *Viscum album***

**(A)** BN-PAGE analysis of mitochondrial samples from *Arabidopsis thaliana* and *Viscum album*. Protein complexes were visualised with Coomassie (left-hand panel). NADH dehydrogenase activities were visualised with NADH and NBT staining (right-hand panel). *Arabidopsis thaliana* mitochondria were purified from callus culture. *Viscum album* mitochondria were Percoll-purified. Complex I, III and V are marked. LPD marks the position of lipoamide dehydrogenase, also known as dihydrolipoamide dehydrogenase, a subunit of the pyruvate dehydrogenase complex.

**(B)** BN-PAGE analysis of crude mitochondrial samples from *Viscum album*. Protein complexes were visualised with Coomassie (left-hand panel). NADH dehydrogenase activities were visualised with NADH and NBT staining (right-hand panel). Position of LPD is marked.



**Figure 4.4: Complex II, III and IV in-gel activity in *Viscum album***

**(A)** BN-PAGE analysis of mitochondrial samples from *Arabidopsis thaliana* and *Viscum album*. Complex II activity was visualised with PMS/Succinate/NBT staining. *Viscum album* samples from crude purification (Left-hand panels) and Percoll purification (Right-hand side).

**(B)** BN-PAGE analysis of crude mitochondrial samples from *Viscum album*. Complex III activity was visualised using ECL reagents and high performance chemiluminescence film detection. Protein loading was visualised by Coomassie staining (Left-hand panel). *Viscum album* mitochondria were from a crude preparation.

**(C)** BN-PAGE analysis of mitochondrial samples from *Arabidopsis thaliana* and *Viscum album*. Complex IV activity was visualised with DAB/cytochrome *c* staining. *Viscum album* mitochondria were from a crude preparation.

blue activity band was seen in both crude and Percoll-purified mitochondria samples just above 146 kDa, indicating the presence of respiratory complex II (Figure 4.4A). Interestingly these were of a higher molecular weight than the *Arabidopsis thaliana* complex II. Respiratory complexes can vary in size between different groups because of different subunit compositions.

In order to detect complex III, *Viscum album* crude mitochondria were separated by BN-PAGE and transferred to PVDF membrane. Visualisation of complex III was achieved by incubation with ECL reagents and detection by high performance chemiluminescence film. This works due to the heme groups bound on complex II, which have peroxidase activity and can therefore be detected by chemiluminescence (Feissner et al. 2003; Sun et al. 2015). A clear band can be seen between the 480 and 720 kDa marker indicating the presence of complex III (Figure 4.4B). This corresponds to a large band seen in a Coomassie stained lane of the same sample. The size is consistent with the size of complex III seen in *Arabidopsis*.

In order to detect complex IV, *Arabidopsis thaliana* and *Viscum album* crude mitochondria were separated by BN-PAGE and activity stained with DAB and cytochrome c (Sabar et al. 2005). A brown precipitate can clearly be seen in the *Arabidopsis* sample at approximately 242 kDa (Figure 4.4C). A fainter band can also be seen in the crude *Viscum album* sample at the same molecular weight, indicating the presence of complex IV.

In summary, these data indicate the presence of complexes II, III and IV in *Viscum album*, supporting the idea that only complex I is absent in *Viscum album* as suggested by (Skippington et al. 2015; Skippington et al. 2017). Complex V could not be assessed due to the potential of chloroplastic ATP synthases influencing the result.

#### **4.2.4 Proteomic analysis indicate that complex I is absent, but the rest of the respiratory chain is intact**

In- gel activity assays indicate that *Viscum album* lacks complex I, but not complex II, III and IV. A proteomic approach was performed as a complementary technique to assess the composition of the respiratory chain.

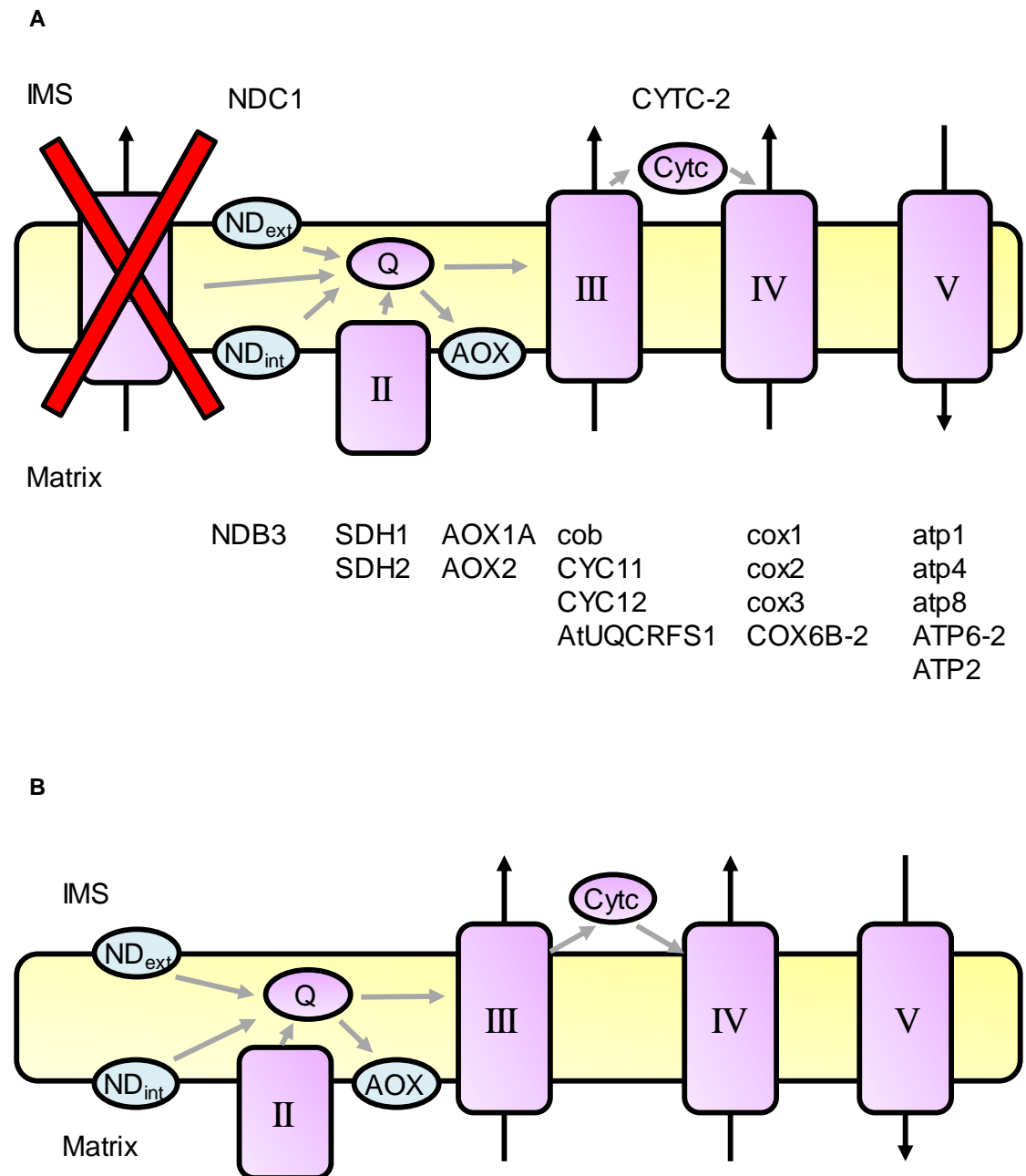
Shotgun proteomics on both crude and Percoll-purified mitochondrial samples was performed. Peptides were digested with trypsin and run through a Thermo Q-Exactive Plus mass spectrometer (All proteomic work was carried out with Dr. Etienne Meyer). When analysing the peptide fragments, there was no *Viscum album* nuclear genome

sequence to compare them, to ascertain their identity. The *Viscum album* nuclear genome has not been sequenced due to its large size, at 201 Gb (Zonneveld 2010). This problem could be overcome by comparing them to a) the *Viscum album* mitochondrial genome and, b) the *Arabidopsis thaliana* nuclear genome. By comparing to the *Viscum album* mitochondrial genome, the presence of mitochondrially encoded proteins could be shown. This as an important control as it confirms that mitochondrial genes are expressed and are not pseudogenised. By comparing to the *Arabidopsis thaliana* nuclear genome, proteins that are highly conserved between the two species can be identified. A caveat is that if a protein is not identified by this method it does not necessarily mean that the protein is absent in *Viscum album*. It could instead be that the sequence of the *Arabidopsis thaliana* and *Viscum album* homologs have diverged to such an extent that *Viscum album* homolog cannot be identified based on the *Arabidopsis thaliana* sequence. Using these two methods numerous mitochondrial proteins were identified in *Viscum album* samples (see appendix), as well as numerous contaminating proteins of chloroplastic and peroxisomal origin.

Using the *Viscum album* mitochondrial genome subunits were identified from complex III, IV and V (Table 4.2; see appendix). Using the *Arabidopsis thaliana* nuclear genome subunits were identified from complex II, III, IV and V (Table 4.3; see appendix). No subunits from complex I were identified. These data indicate that the mitochondrially encoded genes are translated and that protein subunits from complex II, III, IV and V, but not complex I, are encoded in the nuclear genome (Figure 4.5A).

Lineages that have lost complex I, for example the yeast *Saccharomyces cerevisiae*, are increasingly reliant on the alternative respiratory pathway, consisting of alternative NADH dehydrogenases, to bypass the loss of complex I, and alternative oxidases, to bypass complex III and IV. The alternative dehydrogenases oxidise NADH, thereby regenerating the NAD<sup>+</sup> pool, but do not translocate protons. Plants contain type II NAD(P)H dehydrogenases, which are alternative dehydrogenases found in plants, fungi, protists and bacteria, but not in animals (Rasmusson et al. 2008). It was hypothesised by Skippington et al. that *Viscum scurruloideum* would be reliant on alternative dehydrogenases to compensate for the lack of complex I. My proteomic results show the presence of two alternative dehydrogenases (NDB3 and NDC1) showing that *Viscum album* is able to make use of the alternative respiratory pathway. Two alternative oxidases, AOX2 and AOX1A, were also found. Alternative oxidases allow the reduction of oxygen without any proton translocation, and have previously been reported to be increased in plant complex I mutants (Fromm et al. 2016). The presence of this alternative respiratory pathway in *Viscum album* mitochondria shows that it retains the





**Figure 4.5: Proteins in *Viscum album* mitochondria detected by mass spectrometry**

**(A)** Schematic of a plant respiratory chain, with proteins identified in *Viscum album*. The red cross indicates that no complex I subunits were detected. The gray and black arrows represent the flow of electrons and protons, respectively. Protein names in lowercase are mitochondrially encoded, protein names in upper case are nuclear encoded.

**(B)** Schematic of a potential respiratory chain in *Viscum album*. The gray and black arrows represent the flow of electrons and protons, respectively. Adapted from James et al. 2013.

**Table 4.2: Proteins identified by comparison to *Viscum album* mtDNA**

Protein group	Proteins identified based on mtDNA
Complex III	cob
Complex IV	cox1 cox2 cox3
Complex V	atp1 atp4 atp8
Ribosomal protein	rps3 rps4 rpl16

**Table 4.3: Proteins identified by comparison to *Arabidopsis thaliana* database**

Protein group	Proteins identified based on <i>A.t.</i> database
Complex II	SDH1 SDH2
Complex III	RIESKE CYC11 CYC12
Complex IV	cox2 COX6B-2
Complex V	ATP-Synthase subunit beta-1 ATP1 ATP6-2

ability to bypass the requirement of complex I and also regenerate the NAD<sup>+</sup> and ubiquinone pool. This would presumably come at the expense of lower ATP production, due to the decrease in proton translocation.

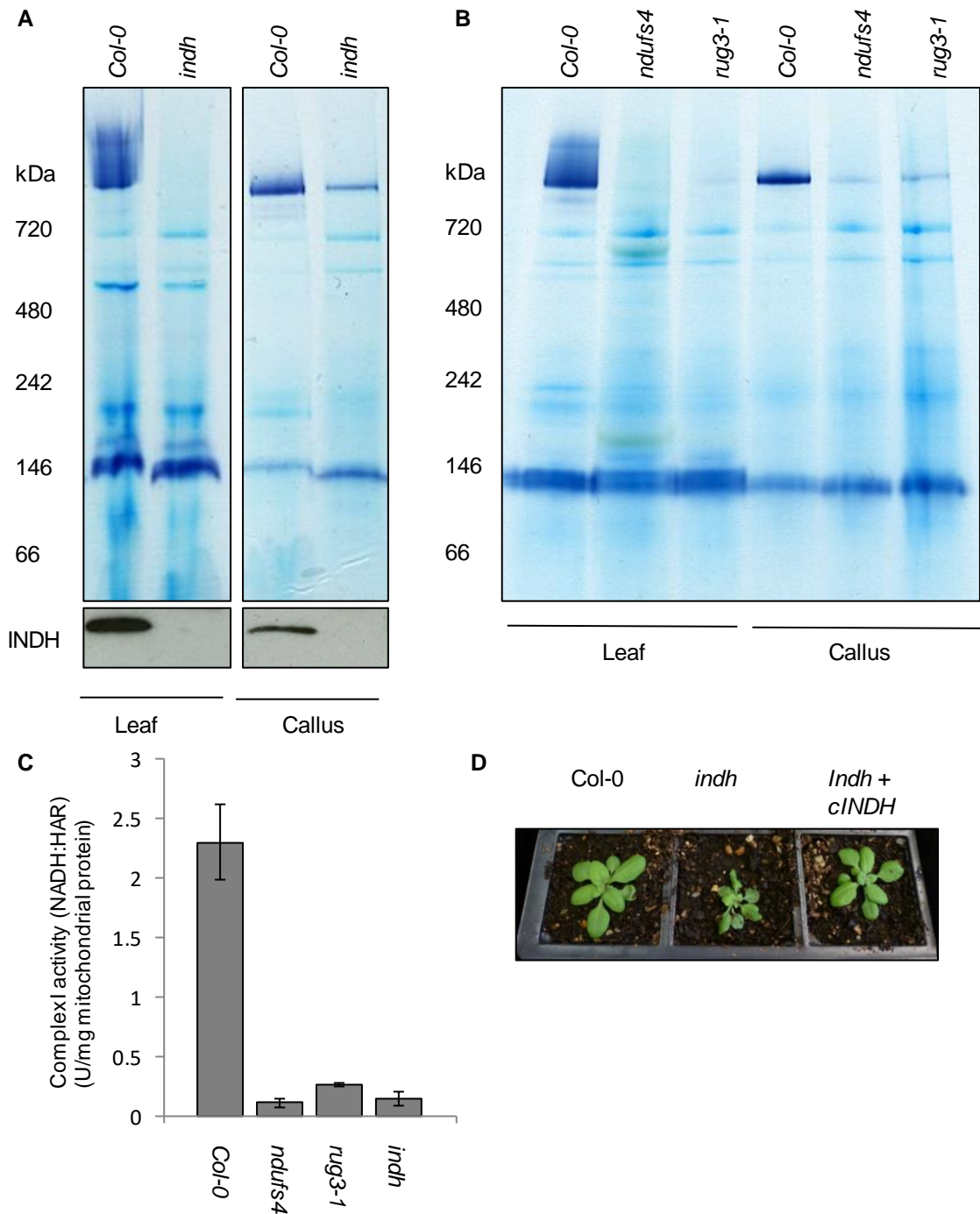
As well as respiratory chain subunits, three ribosomal proteins encoded by the mitochondrial genome were identified, rps3, rps4 and rpl16 (Table 4.2). These genes were predicted have been lost by Petersen et al 2015, but present by Skippington et al. 2017. Their proteomic identification shows they have not become pseudogenes, and also that *Viscum album* retains the machinery for mitochondrial translation.

Overall the proteomic data from *Viscum album* supports the findings of the in-gel activity assays. Subunits were identified from all respiratory complexes, including the alternative pathways, except from complex I. This allows a reconstruction of the *Viscum album* respiratory chain (Figure 4.5B).

#### **4.2.5 Analysis and comparison of *Arabidopsis thaliana* complex I mutants**

*Viscum album* appears unique, as a multicellular organism, in having lost complex I. In humans, slight defects in complex I are often fatal. However, it has previously been seen that *Arabidopsis* complex I mutants can have relatively mild defects during vegetative growth despite having problems during germination (Wydro et al. 2013, Figure 4.6D). Studying how *Arabidopsis thaliana* mutants survive with low complex I levels, may give clues as to how *Viscum album* survives with no complex I.

Previous BN-PAGE analysis of a mutant in the complex I assembly factor, INDH, suggests that this mutant completely lacks complex I (Wydro et al. 2013). However, mutants in the yeast homolog display residual complex I activity (Bych et al. 2008; chapter 3). In order to ascertain if complex I levels are absent in *Arabidopsis indh*, BN-PAGE analysis was carried out on mitochondria isolated from both leaves and callus of *indh* mutants. Callus tissue is equivalent to undifferentiated, non-photosynthetic, meristematic tissue. In leaf-derived mitochondria no complex I activity stain was observed in the *indh* mutants (Figure 4.6A), in agreement with the findings of Wydro et al. 2013. However, low levels of complex I activity stain were seen in callus-derived *indh* mitochondria, indicating that this line does retain the ability to assemble low levels of mature complex I (Figure 4.6A). These data indicate *indh* plants have undetectable levels of complex I in leaf tissue, but do have low levels of complex I present in meristem-like tissues. This suggests that certain tissues are able to survive with extremely low, or absent, levels of complex I.



**Figure 4.6: Comparison of complex I in *Arabidopsis thaliana* mutants**

**(A)** BN-PAGE analysis of mitochondrial samples from *Arabidopsis thaliana* Col-0 and *indh* purified from leaf or callus culture. NADH dehydrogenase activities were visualised with NADH and NBT staining

**(B)** BN-PAGE analysis of mitochondrial samples from *Arabidopsis thaliana* Col-0, *ndufs4* and *rug3-1* purified from leaf or callus culture. NADH dehydrogenase activities were visualised with NADH and NBT staining.

**(C)** NADH:HAR oxidoreductase activity in leaf mitochondria from *Arabidopsis thaliana* Col-0, *ndufs4*, *rug3-1* and *indh*. Error bars represent s.d.

**(D)** Growth of 3-week-old Col-0, *indh* and *indh* + *cINDH*.

In order to see if this tissue dependent pattern of complex I levels was unique to *indh*, or is present in different complex I mutants, BN-PAGE analysis was carried out on mitochondria isolated from both leaves and callus of two other complex I mutants: *ndufs4* and *rug3-1*. *NDUFS4* encodes a structural subunit of complex I, the 18 kDa subunit, and has previously been studied in (Meyer et al. 2009). *RUG3* encode an RNA editing factor of *nad2*, and has previously been shown to have greatly decreased complex I levels (Kühn et al. 2011). BN-PAGE analysis on leaf-derived mitochondria show that *ndufs4* and *rug3-1* have very low levels of complex I (Figure 4.6B). Callus-derived mitochondria have higher levels of complex I. These data show that, similarly to *indh*, complex I mutants have very low levels of complex I in the leaves, but higher levels in meristem-like tissue.

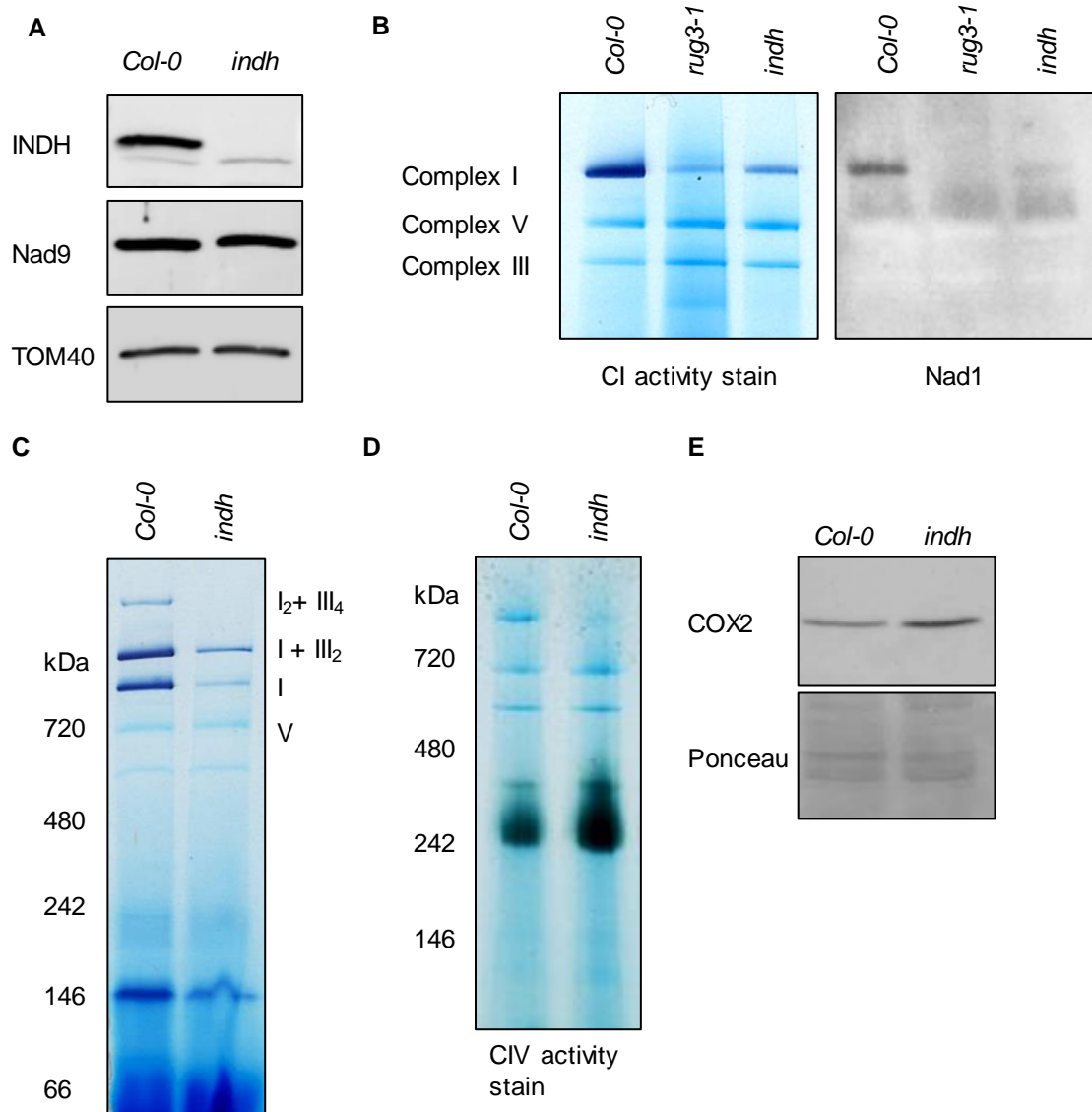
In order to see if *Arabidopsis* complex I mutants *indh*, *rug3* and *ndufs4* completely lack complex I in the leaves, a spectrophotometric assay measuring electron transfer from NADH to the artificial electron acceptor HAR was carried out. These are more sensitive than BN-PAGE activity assays and so will be able to detect complex I activity unobservable in BN-PAGE. Low levels, between 5 and 12% of WT, of complex I activity were seen in all three mutants (Figure 4.6C), indicating residual complex I activity in leaf tissue.

Taken together these data show that tissues are able to survive with only residual levels of complex I in *Arabidopsis thaliana*.

### 4.2.6 Further characterisation of *Arabidopsis thaliana indh* mutant

The role of INDH in complex I assembly was previously studied by Wydro et al. 2013. However, the knock-out mutant was obtained late in the study and most of the experiments were undertaken with a complemented line, B10, which initially had low levels of INDH protein, but gradually over the course of development increased to high levels (~250% in mature rosette leaves). Therefore, further characterisation of the *indh* knock-out mutant was desirable.

The study found that as well as being defective in complex I assembly, *indh* mutants had a defect in mitochondrial translation. *In organello* translation assays revealed that total mitochondrial translation was substantially decreased. In addition, the protein levels of the mitochondrially encoded complex I subunit, *nad9*, was greatly decreased. The mitochondrially encoded complex IV subunit *cox2* was decreased by 50%. In order to see if these findings were also seen in the knock out *indh* mutant, Western blot analysis was carried out to detect the protein levels of *nad9* and *cox 2*. *Nad9* protein levels were



**Figure 4.7: Further characterisation of *Arabidopsis thaliana indh* mutant**

**(A)** Levels of the mitochondrial proteins INDH and Nad9 in wild type (Col-0) and *indh* mutant. TOM40 was used as a marker protein for equal transfer and loading. Mitochondria were purified from callus culture and analysed by immunoblotting with specific antibodies.

**(B)** Complex I and Nad1 protein levels in *Arabidopsis thaliana* Col-0, *rug3* and *indh*. Respiratory complexes were separated by BN-PAGE and stained with NADH/NBT (left-hand panel) or transferred to PVDF and immunolabelled with an antibody specific to Nad1.

**(C)** Respiratory supercomplexes in *Arabidopsis thaliana* Col-0 and *indh*. NADH dehydrogenase activities were visualised with NADH and NBT staining. Mitochondria were purified from callus culture and respiratory complexes solubilised with digitonin.

**(D)** Complex IV levels in *Arabidopsis thaliana* Col-0 and *indh*, visualised by BN-PAGE followed by Complex IV activity staining.

**(E)** Levels of the mitochondrial proteins COX2 in *Arabidopsis thaliana* Col-0 and *indh* mitochondria. Mitochondria were purified from callus culture. Ponceau S stain was used as a control for equal transfer and loading.

unaffected in the *indh* knock-out mutant (Figure 4.7A). Cox2 protein levels were increased, rather than decreased, in the *indh* mutant (Figure 4.7E). BN-PAGE followed by Western blot analysis was performed to detect the presence of another mitochondrially encoded complex I subunit, nad1, in mature complex I. Very low levels of nad1 protein was observed in *indh* at a position corresponding to complex I (Figure 4.7B), suggesting mitochondrial translation has not completely stopped. These results suggest that the role of INDH in mitochondrial translation is not as simple as previously thought and that further investigation is needed.

Respiratory complexes can form ordered associations with one another and arrange themselves into larger complexes, often called supercomplexes. Supercomplexes containing complex I are not easily detectable by BN-PAGE when samples are prepared with DDM as the solubilising agent. However, if digitonin, one of the mildest detergents, is used instead respiratory supercomplexes can be separated (Wittig et al. 2006). The following respiratory super-complexes containing complex I can be visualised in *Arabidopsis*: complex I and a dimer of complex III ( $I + III_2$ ) and a dimer of complex I and a tetramer of complex III ( $I_2 + III_4$ ) (Eubel et al. 2003). In order to study the role of INDH in supercomplex formation mitochondria from Col-0 and *indh* were solubilised with digitonin and separated by BN-PAGE. In the *indh* mutant no  $I_2 + III_4$  can be detected, while low levels of  $I + III_2$  and monomeric complex I can be (Figure 4.7C). The levels of  $I + III_2$  are greater than monomeric complex I in the *indh* mutant. These data show that *indh* is unable to form  $I_2 + III_4$ , presumably limited by the low levels of complex I.

It has previously been shown that in complex I mutants, other respiratory complexes can be upregulated to compensate for the decrease in proton pumping from complex I (Fromm et al. 2016; Hsu et al. 2014). In order to determine if the other respiratory complexes are upregulated in the *Arabidopsis indh* mutant, BN-PAGE and in-gel activity assays were undertaken. No repeatable difference was seen when a complex II activity stain was performed. However, a clear increase in complex IV activity stain is seen in *indh* mitochondria compared to Col-0 (Figure 4.7D). In order to confirm the increase in complex IV, Western blot analysis with an antibody against the complex IV subunit cox2 was performed. An increase in cox2 subunit was observed in *indh* (Figure 4.7E), consistent with previous observations in complex I mutants (Hsu et al. 2014). These data suggest that complex IV is upregulated in the *indh* mutant.

### 4.3 Discussion

Previous mitochondrial genome studies by two groups suggested that members of the genus *Viscum* had lost complex I, due to an absence of *nad* genes in the mitochondrial genome. It remained possible that the whole suite of mitochondrial *nad* genes had been transferred to the nuclear genome. Nuclear transfer of mitochondrial genes has previously been shown in angiosperms, for example the transfer of complex IV subunit *cox2* in Legumes (Nugent & Palmer 1991). Over evolution *nad* gene content of the mitochondrial genome has changed, for example humans encode 7 *nad* genes, whereas plants encode 9. However, the authors in Skippington et al. 2015 argue that the mechanistic challenges in transferring all nine genes which code hydrophobic products would make the likelihood of this vanishingly small. Also, the presence of *nad* pseudogenes in *Viscum album* argues against nuclear transfer (Skippington et al. 2017). Therefore, they argue, the absence of *nad* genes in the mitogenome of *Viscum* represents a loss of complex I. Nevertheless, despite the very low probability of nuclear transfer, the possibility remains. It was therefore necessary to test if complex I was truly absent. I elected to do this through a biochemical and proteomic approach. By obtaining purified mitochondria and “looking” inside I would avoid the ambiguities which come from working with sequences. No complex I was detected in *Viscum album* through in-gel activity assays, or by proteomics. These data strongly support the conclusions from the mitochondrial genomic studies. However, it still remains possible, although extremely unlikely, that very low levels of complex I exist which are below the detection limit for the assays I used. This would require that, as well as all *nad* genes being transferred to the nucleus, all 40+ genes which encode complex I subunits have diverged from other species to such a degree, or be at such low levels, that they were not recognisable by proteomics using the *Arabidopsis* database as a template. The only way to say for certain if complex I is gone would be to sequence the genome of *Viscum album*. This has not been undertaken so far due to the prohibitively large nuclear genome size, at 201 Gb (Zonneveld 2010), although may be undertaken as part of future work. The assumption for the remainder of this discussion will be that complex I has indeed been lost.

The loss of complex I is unprecedented in multicellular eukaryotes. All but four lineages of aerobic eukaryotes contain complex I. The exceptions are the Cryptomycotan *Rozella allomyces* (James et al. 2013); the lineage which contains dinoflagellates, apicomplexans, *Chromera* and *Oxyrrhis* (Flegonotov et al. 2015) and two groups of yeast (Gabaldón et al. 2005). All of these are unicellular, making the loss in *Viscum*, which is



multicellular with a complex body-plan, all the more striking. Across the eukaryotes major changes to the respiratory chain are infrequent and almost always involving anaerobes. Of these most examples are parasites, illustrating again that the selection pressures in evolving a parasitic lifestyle can lead to organellar genome reduction.

Despite agreement that *Viscum* had lost complex I, the two 2015 studies disagreed on the presence of other mitochondrial genes. Petersen, et al. suggested that all mitochondrial respiratory genes, as well as most of the other ones, in *Viscum album* were non-functional, or pseudogenised. Although two different species were studied it would seem unlikely that they possess radically different gene complements. The authors of the first study later reanalysed the *Viscum album* sequence data and vigorously disagreed with the conclusions from Petersen et al (Skippington et al. 2017). Here, I was able to solve this disagreement. Our data strongly support the findings from Skippington et al 2017, that only respiratory complex I is missing. By detecting mitochondrially encoded proteins by proteomics, we demonstrated that the genes are expressed. By detecting the respiratory complexes with in-gel assays we demonstrated that the proteins are functional. Therefore, despite lacking complex I, the rest of the respiratory chain remains intact. Although subsequent work by Dr. Etienne Meyer, who collaborated on this project, suggest the respiratory chain is down-regulated. However, in this case the remodelling of the respiratory chain in *Viscum* is not as severe as that seen in *Chromera velia*, a phototrophic relative of parasitic apicomplexans, which has lost both complex I and III (Flegonotov et al. 2015). In *Chromera velia*, alternative dehydrogenases are used to bypass the requirement for complex I, and putative lactate:cytochrome c oxidoreductases are predicted to bypass the requirement for complex III, by transferring electrons from lactate to complex IV.

As the first studied example of a multicellular eukaryote that has lost complex I it is likely that *Viscum* has evolved unique features to deal with this metabolic upheaval. For example, extensive metabolic rearrangement is to be expected in order to compensate for the substantial decrease in mitochondrial ATP production. This suggests a model whereby *Viscum* generates only small, or no, mitochondrial ATP. The proton gradient is maintained by complexes II-IV, but is used instead to power transport. ATP could be derived instead from glycolysis, or from the host. The photosynthetic rate is known to be low in *Viscum album* (Zuber 2004), making it likely that mistletoe is highly dependent on host nutrition, and that photosynthetically derived ATP is unlikely to fully compensate for the decreased levels of mitochondrial ATP. The chloroplast genome of *Viscum* has also been sequenced, and displays signs of reductive evolution, possibly signifying increasing nutritional dependence of the host (Petersen et al. 2015b), however it retains its

photosynthetic machinery. Interestingly, similar to the mitochondrial genome, the *ndh* genes, which encode subunits for the chloroplast NAD(P)H dehydrogenase complex, have been lost (Petersen et al. 2015b). Future work, in progress with Dr. Etienne Meyer, aims to characterise the metabolic rearrangements seen in *Viscum album*, to understand how mitochondrial metabolism has adapted to the loss of complex I.

Comparisons to *Arabidopsis* mutants that have a low level of complex I can also be instructive. Results show that leaf tissue can survive with trace amount of complex I (Figure 4.6), and that complex IV is upregulated (Figure 4.7D). A recent proteomic study, (Fromm et al. 2016), characterised a complex I mutant, *ca1ca2*, in order to understand *Arabidopsis*'s response to complex I absence. The study found an upregulation of complex II, IV and alternative oxidases, as well as extensive metabolic change and reorganisation. For example, certain parts of the TCA cycle, glycolysis, fermentation, amino acid catabolism and stress responses were increased, but photosynthetic metabolism was decreased. It is probable that many changes seen in the *Arabidopsis ca1ca2* mutant will also be seen in *Viscum album*. However, mutant physiology is by definition aberrant, and care should therefore be taken in applying these observations to *Viscum*, which has had millions of years of evolution to adapt to the changes brought about by a lack of complex I.

*Viscum* sp. are currently the only parasitic plants known to have lost complex I. However, the mitochondria of parasitic plants remain relatively understudied. Apart from *Viscum*, only *Rafflesia lagascae* (Molina et al. 2014) and *Rafflesia cantleyi*, *Rafflesia tuan-mudae* and *Sapria himalayana* (Xi et al. 2013) have been studied, meaning there is a great deal left to be investigated. These mitochondrial genomes appeared normal in terms of gene content. It is likely that the same selective pressures that led to *Viscum* sp. losing complex I have been replicated elsewhere during the 13 independent origins of plant parasitism. If true, there exist other parasitic plants with non-canonical respiratory chains waiting to be discovered. To identify these, targeted sequencing of mitochondrial genomes of representative species from every lineage that evolved parasitism could be carried out, followed by nuclear genome sequencing, or RNAseq. Further biochemical and proteomic work, like that carried out in this study, could then be performed to characterise these respiratory chains. Such studies would give insight into respiratory chain evolution and open up fundamental questions about the requirement of respiratory complexes for multicellular life.

## 5 Functional characterization of FeS assembly proteins in *Arabidopsis thaliana* mitochondria

### 5.1 Introduction

The FeS clusters present in respiratory complex I, II and III are assembled and inserted by the mitochondrial ISC pathway. The ISC pathway has been most intensively studied in the yeast model organism, *Saccharomyces cerevisiae*, or Baker's yeast. However, complex I has been lost in Baker's yeast making the effect of FeS assembly defects on complex I hard to study. Studies in humans have allowed the connection between FeS assembly and respiratory complexes to be explored. Patients with defects in components of the ISC pathway have been found with respiratory chain defects. However, studies in humans are limited by the availability of patient mutations or rely on cell culture and RNAi studies, which may have limited relevance to the physiology of a multi-cellular organism.

Plants provide a good alternative system to study the impact of FeS assembly defects on complex I and complex II. They are complex multicellular eukaryotes with extensive genetic resources, for example a good selection of T-DNA knock-out mutants. With the exception of the FeS assembly factor INDH (Wydro et al. 2013) the connection between FeS assembly and respiratory complex I and II has never been previously studied. In this chapter I will study four different genes which encode putative FeS assembly proteins, ISU1, NFU4, NFU5 and GRXS15 (Table 5.1), to ascertain the role they play in FeS cluster provision for the respiratory chain. Additionally, I will try to assign their role in FeS assembly as there is currently little agreement about which FeS carrier proteins affect which subset of Fe-S requiring enzymes (Figure 5.1A).

ISU1, the plant homolog of yeast Isu1 and human ISCU (Table 5.1), is an important scaffold for FeS assembly and appears to have a conserved role across eukaryotes. Three *ISU* paralogs exist in *Arabidopsis*: *ISU1*, *ISU2* and *ISU3*. Through GFP localisation studies ISU1, ISU2 and ISU3 have been shown to be mitochondrially targeted (Léon et al. 2005). *ISU1* is thought to be the main scaffold, due to its high expression levels relative to the other two paralogs (Léon et al. 2005). *In vitro* studies have shown ISU1 is able to bind an [2Fe-2S] cluster (Frazzon et al. 2007). ISU1 has been shown to interact with the cochaperone protein HSCB, and together can stimulate the ATPase activity of the chaperone protein HSCA1 (Xu et al. 2009). The same study also indicated that ISU1 can be retained in the cytosol, as well as mitochondria, through interaction with HSCB,

similarly to the cytosolic ISC proteins seen in the mammalian system. Taken together, these data suggest that ISU1 plays a similar role in plants as it does in other eukaryotes. However, *ISU1* has been minimally characterised in *Arabidopsis*.

NFU (NifU-like) proteins are named after the NifU gene in the nitrogen fixation (*nif*) regulon in the nitrogen fixing bacteria *Azotobacter vinelandii*. Along with the cysteine desulphurase NifS, NifU is required for the maturation of FeS clusters for the nitrogen fixing enzyme, nitrogenase (Yuvaniyama et al. 2000). NifU is a modular protein with three parts (Agar et al. 2000). The N-terminal section contains three conserved cysteines, and is homologous to the FeS scaffold ISU. The central section has a ferredoxin-like domain, with four cysteine residues which coordinate a permanent FeS cluster (Agar et al. 2000). The C-terminal domain is homologous to mitochondrial NFU proteins, with 2 conserved cysteines. Therefore, the 3 domain *Azotobacter vinelandii* NifU protein is a fusion of 3 proteins that are usually separate in bacteria, and always separate in eukaryotes, as *ISU*, *FDX* and *NFU* (Figure 5.1B). The function of NFU proteins has been studied in bacteria, yeast and humans and have been shown to be important for mitochondrial FeS cluster maturation. They have also been studied in plant chloroplasts and been shown to be important in the SUF FeS assembly pathway (Touraine et al. 2004).

*Saccharomyces cerevisiae* contain just one NFU gene, *Nfu1*, which encodes a protein targeted to the mitochondria. Studies show that *nfu1Δ* strains are defective in the FeS containing enzymes aconitase (~20-70% of WT activity) and respiratory complex II (~30-50% of WT activity) (Schilke et al. 1999; Navarro-Sastre et al. 2011; Uzarska et al. 2016; Melber et al. 2016). Additionally, *nfu1Δ* is defective in lipoate cofactor biosynthesis, possibly due to the requirement of a FeS cluster in the lipoate cofactor biosynthetic enzyme, lipoate synthase. Consequently, lipoate containing enzymes are severely affected, for example PDH (~20% of WT activity) and KGDH (~20-40% of WT activity) (Navarro-Sastre et al. 2011; Melber et al. 2016; Uzarska et al. 2016). When *Nfu1* is combined as a double mutant with *Isu1* in *Saccharomyces* the defects are more pronounced and growth is severely impaired (Schilke et al. 1999). The phenotypes of *nfu1Δisu1Δ* strains and the modular structure of NifU highlight the importance and functional closeness of both *Nfu1* and *Isu1* for FeS cluster maturation. Further work in the last decade has suggested NFU proteins act late in the FeS assembly pathway and are specifically involved in [4Fe-4S] cluster maturation (Melber et al. 2016). Recently several cases of mitochondrial disease, usually fatal, have been linked to mutations in the *NFU1* gene (Navarro-Sastre et al. 2011; Cameron et al. 2011; Ferrer-Cortès et al. 2013; Invernizzi et al. 2014; Nizon et al. 2014; Tonduti et al. 2015; Ahting et al. 2015;

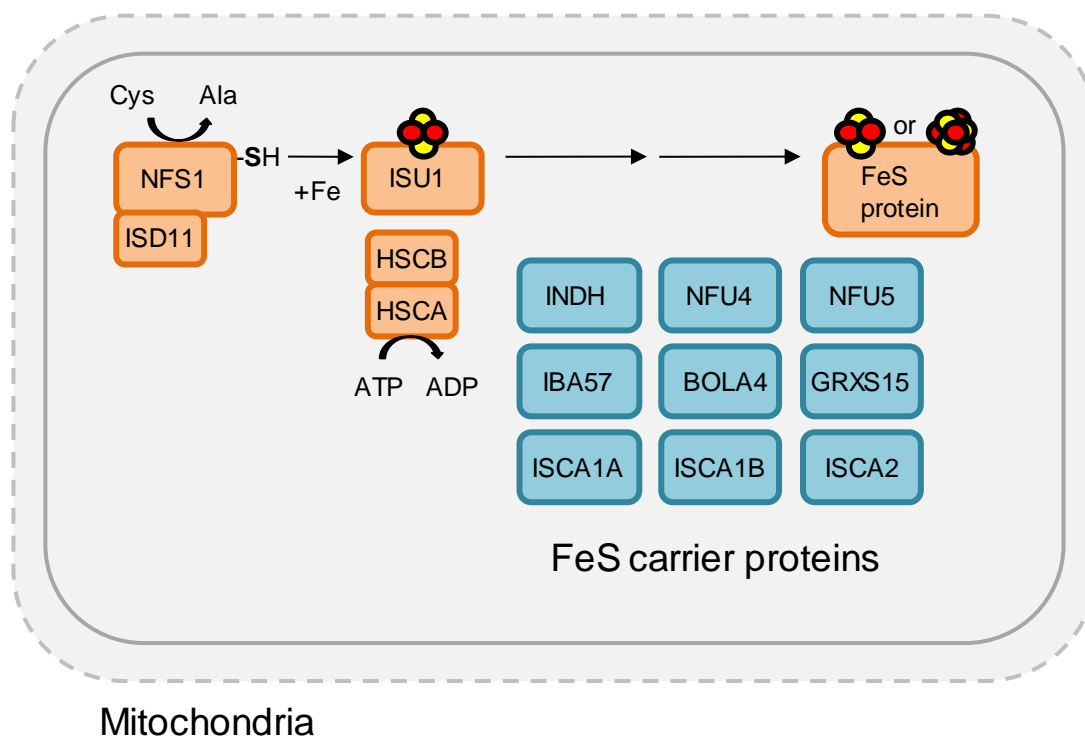
Tort et al. 2016; Lebigot et al. 2017). Most *NFU1* patients display infantile encephalopathy or pulmonary hypertension, with symptoms such as hyperglycinemia and lactic acidosis. Biochemical tests showed decreases in certain FeS dependent enzymes such as the respiratory complexes I,II and III as well as PDH. Death often occurs within the first fifteen months. This is in striking contrast to the relatively mild phenotype of yeast *nfu1Δ* mutants.

The majority of eukaryotes have only one *NFU* gene, the product of which is targeted to the mitochondria. However, as previously reported, *Arabidopsis thaliana* has five genes with homology to *NFU* (Léon et al. 2003). *NFU1*, *NFU2* and *NFU3* encode proteins which are targeted to the chloroplast, and have high sequence similarity to the cyanobacterial homologs from *Synechocystis* PCC6803. *NFU2* has been shown to be important for [4Fe-4S] and [2Fe-2S] ferredoxin cluster assembly in the chloroplasts (Touraine et al. 2004). *NFU4* (AT3G20970) and *NFU5* (AT1G51390) encode proteins which are predicted to be targeted to the mitochondria, which has been experimentally verified for *NFU4*, but not *NFU5* (Léon et al. 2003), and have high sequence similarity to other eukaryotic mitochondrial NFUs for example to human and yeast *NFU1*. The physiological role of *NFU4* and *NFU5* in *Arabidopsis* has never been previously investigated.

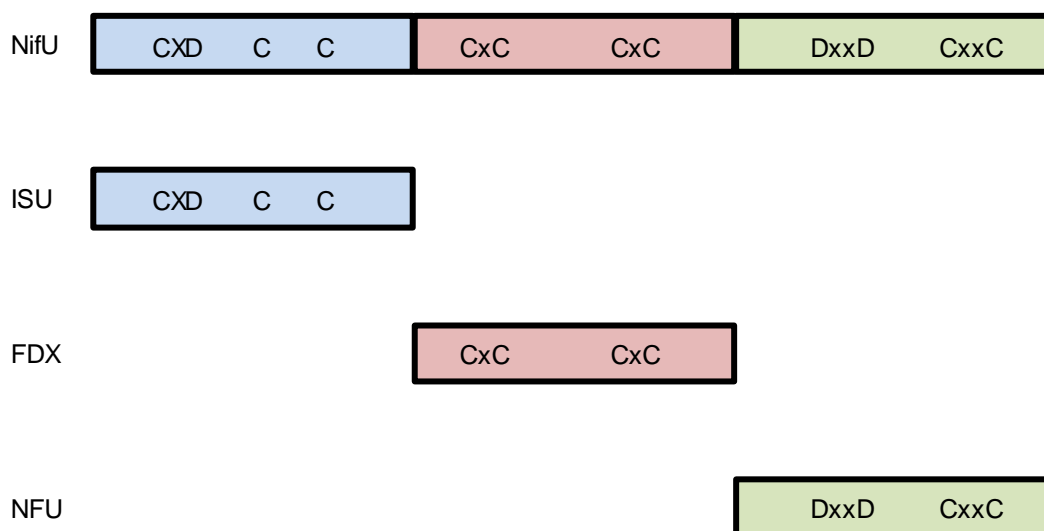
Glutaredoxins (Grxs) are small proteins that act as redox enzymes, usually as oxidoreductases (Ströher & Millar 2012). They classically function to deglutathionylate proteins, thereby removing a modification that protects sensitive thiols during oxidative stress. Certain Grxs were also hypothesised to have a role in FeS assembly. A study in 2002 suggested *Grx5* in *Saccharomyces cerevisiae* was required for mitochondrial FeS assembly (Rodríguez-Manzanique et al. 2002). Additionally, it was shown that some Grxs could bind FeS clusters (Lillig et al. 2005). Since then numerous other studies supported the involvement of Grxs in FeS assembly. Grx proteins now commonly appear in FeS assembly pathway models in both the mitochondria, cytosol and chloroplast. Sometimes they enjoy a central role: for example, directly receiving the nascent FeS cluster from *Isu1*, and passing it to other downstream carriers (Uzarska et al. 2013; Couturier et al. 2013; Stehling et al. 2014; Braymer & Lill 2017). In other models, they take a more peripheral role, being one of many potential FeS carriers (Balk & Schaedler 2014; Maio & Rouault 2015; Rouault & Maio 2017).

The *Arabidopsis thaliana* homolog of yeast *Grx5*, is *GRXS15*. It is one of 31 Grxs encoded by the *Arabidopsis thaliana* genome. The localisation of *GRXS15* was uncertain

A



B



**Figure 5.1: Mitochondrial ISC pathway in *Arabidopsis thaliana***

**(A)** Outline of the mitochondrial ISC pathway in *Arabidopsis thaliana*. FeS carrier proteins are highlighted in blue, other ISC pathway components are highlighted in orange. FeS clusters are shown as red and yellow circles. Mitochondrial inner membrane is indicated with a solid grey line, the mitochondrial out membrane is indicated with a dashed grey line.

**(B)** Schematic of the modular protein NifU from *Azotobacter vinelandii* as well as the eukaryotic proteins ISU, FDX and NFU (adapted from Benz et al. 2016). The N-terminal section of NifU is homologous to the eukaryotic ISU proteins (blue boxes) and contains three conserved cysteines which coordinate a transient [2Fe-2S] cluster. The central section is homologous to the eukaryotic FDX proteins (red boxes) and contains four conserved cysteines which coordinate a permanent [2Fe-2S] cluster. The C-terminal section of NifU is homologous to the eukaryotic NFU proteins (green boxes) and contains two conserved cysteines which coordinate a transient [2Fe-2S] cluster. Additionally, conserved aspartate residues have been proposed to aid the transfer of newly formed FeS clusters.

**Table 5.1: Mitochondrial ISC pathway in *Arabidopsis thaliana***

Name	Gene id	Homologs (yeast/human)	Proposed function
NFS1	AT5G65720	Nfs1/NFS1	Cysteine desulfurase
ISD11	AT5G61220	Isd11/ISD11	Helper protein for NFS1
ISU1	AT4G22220	Isu1/ISCU	FeS scaffold
ADXR	AT4G32360	Arh1/FDXR	Electron donation
ADX1	AT4G21090	Yah1/FDX2	Electron donation
ADX2	AT4G05450	Yah1/FDX2	Electron donation
HSCA1	AT4G37910	Ssq1/HSPA9	Chaperone
HSCA2	AT5G09590	Ssq1/HSPA9	Chaperone
HSCB	AT5G06410	Jac1/HSC20	Chaperone
NFU4	AT3G20970	Nfu1/NFU1	FeS carrier
NFU5	AT1G51390	Nfu1/NFU1	FeS carrier
ISCA1A	AT2G16710	Isa1/ISCA1	FeS carrier
ISCA1B	AT5G03905	Isa1/ISCA1	FeS carrier
ISCA2	AT2G36260	Isa2/ISCA2	FeS carrier
GRXS15	AT3G15660	Grx5/GLRX5	FeS carrier/ redox reactions
BOLA4	AT5G17560	Aim1/BOLA3	FeS carrier/ interact with Grx
INDH	AT4G19540	Ind1/NUBPL	FeS carrier
IBA57	AT4G12130	Iba57/IBA57	FeS carrier

for a long time, with some reports placing it in the mitochondria (Bandyopadhyay et al. 2008; Klodmann et al. 2011; Salvato et al. 2014), others in the chloroplast (Cheng 2008; Huang et al. 2013) and one in both (Couturier et al. 2014). Two recent papers (Moseler et al. 2015; Stroehrer et al. 2015) conclude that GRXS15 is localized in the mitochondrial matrix. Moseler et al. show that GRXS15 protein is able to coordinate and transfer a [2Fe-2S] cluster *in vitro* and that mutation of the conserved Lys83 led to a ~50% decreased activity of aconitase, an enzyme that requires a FeS cluster. Based on this the authors concluded that “GRXS15 is essential for iron-sulfur protein maturation”. Stroehrer et al. also showed the ability of GRXS15 to bind a [2Fe-2S] cluster and that knock-down mutants had defects in certain lipoate containing enzymes, but not complex I or aconitase. The lack of defect in aconitase is surprising as it is the opposite of what is observed in the Moseler et al. study. Additionally, they report an increased sensitivity to arsenic. However, despite the strong *in vitro* evidence that GRXS15 coordinates an FeS cluster, and the severe growth phenotype observed for both the K83A and knock-down mutants, the FeS defects are surprisingly mild for a protein of central importance in FeS assembly. This suggests that GRXS15 may not play as central a role in FeS assembly as is currently thought.

## 5.2 Results

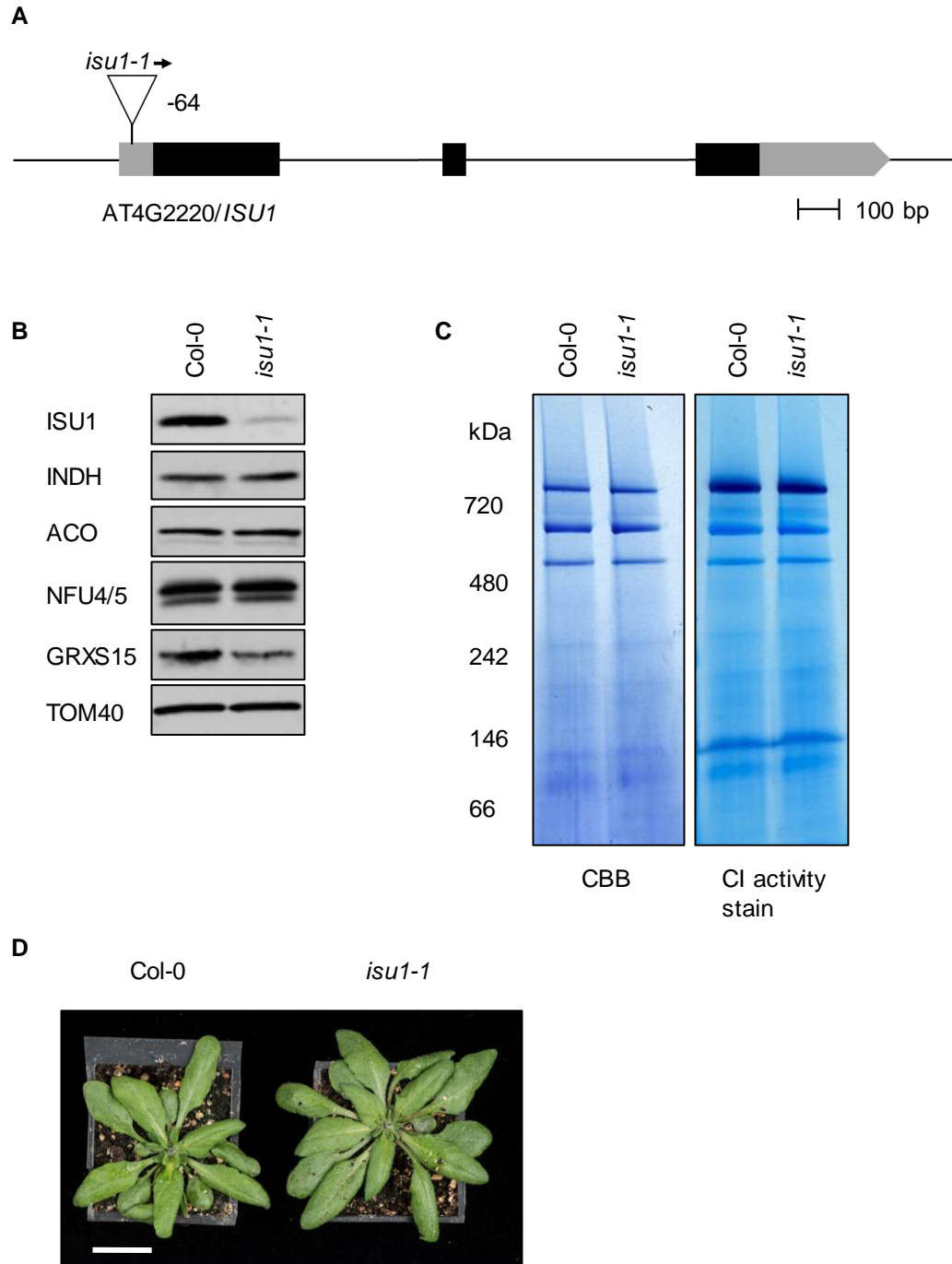
### 5.2.1 Characterisation of the *isu1-1* mutant

In order to study the role of *ISU1* a T-DNA insertion mutant was investigated for growth and biochemical defects.

The *isu1-1* T-DNA insertion line (SALK\_006332) (Table 5.2) was previously obtained from the Nottingham Arabidopsis stock centre. Homozygous plants isolated were by genotyping PCR analysis and initial analysis carried out (Unpublished data: Delphine Bernard, Jo Harrison, Inga Kruse, Janneke Balk). The exact insertion site of the T-DNA was determined by sequencing (Figure 5.2A) to be 64 bp upstream of the ATG sequence, in the 5' UTR. Western blot analysis using an antibody against ISU1 showed a significant decrease in protein levels (Figure 5.2B). This was quantified to be approximately 18% of WT levels.

A previous study in cucumber, *Cucumis sativus*, showed that iron deficiency led to a dramatic decrease in complex I activity and assembly (Vigani et al. 2009). Therefore, it is reasonable to assume that defects in the FeS pathway may lead to a decrease in abundance of complex I, II and III as well a decrease in activity. The abundance of respiratory complexes in *isu1-1* was analysed using BN-PAGE and coomassie staining





**Figure 5.2: Analysis of *isu1-1* mutant**

**(A)** Gene model of *ISU1* and the positions of the T-DNA insertion. Black bars represent exons, grey bars are the 5' and 3' untranslated regions of the transcript. Triangle represent T-DNA insertion, its orientation is marked with an arrow to indicate the outward facing left border primer LBb1.3.

**(B)** SDS-PAGE followed by Immunolabelling of mitochondrial FeS proteins in purified mitochondria samples isolated from callus lines.

**(C)** Respiratory complexes I, III and V separated by BN-PAGE and visualised with coomassie staining (left-hand panel) and by complex I NADH/NBT activity staining (right-hand panel) in wild-type (Col-0) and *isu1-1* mitochondria samples isolated from callus lines.

**(D)** Growth phenotype of 4-week old *isu1-1* T-DNA insertion line and corresponding WT, Col-0. Scale bar = 2 cm.

(Figure 5.2C). No decreases of any respiratory complexes were observed. Activity staining with NBT/NADH, to visualise complex I levels, again showed no decreases. Western blot analysis on isolated mitochondria showed no effect on protein levels of FeS binding proteins aconitase, NFU4, NFU5, INDH and GRXS15 (Figure 5.2B). TOM40 was used as a control for equal loading and transfer.

In order to evaluate the importance of *ISU1* on growth, the *isu1-1* mutant was grown on soil under long-day conditions (Figure 5.2D). No growth phenotype was observed.

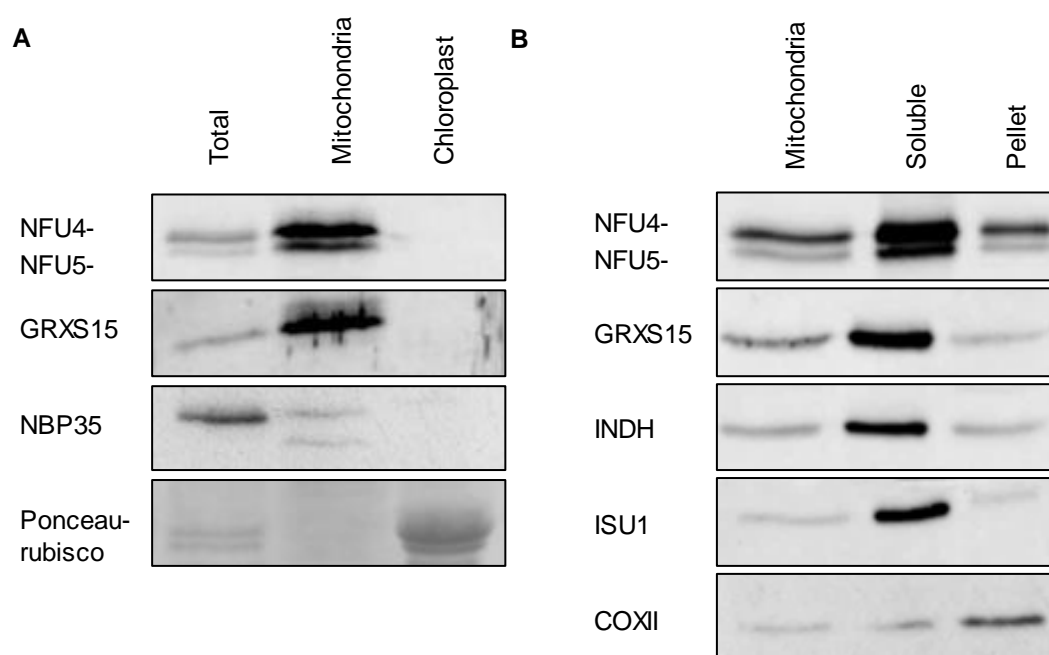
Taken together these data show that, despite an ~80% decrease in ISU1 protein levels, the *isu1-1* mutant does not have a clear phenotype.

### 5.2.2 NFU4 and NFU5 are localised to the mitochondrial matrix

*Arabidopsis thaliana* contains five *NFU* genes. Two of them, *NFU4* and *NFU5*, are predicted to encode proteins that are targeted to the mitochondria. Through GFP localisation studies NFU4 has been shown to be mitochondrially targeted (Léon et al. 2003). However, the localisation of NFU5 has yet to be confirmed. To confirm the localisation of NFU5 Western blot analysis was performed on total protein, mitochondrial and chloroplastic fractions using antibodies that bind both NFU4 and NFU5 (Figure 5.3A). Both NFU4 and NFU5 were enriched in the mitochondrial fraction, similarly to GRXS15 which has previously been shown to be mitochondrially localised. The cytosolic marker NBP35 was enriched in the total protein fraction. No NFU4 or NFU5 was found in the chloroplastic fraction, identified by the enrichment of Rubisco on the ponceau stain.

In order to elucidate the mitochondrial sub-compartment where NFU4 and NFU5 reside, mitochondria were separated by freeze-thawing into a soluble matrix fraction and insoluble pellet fraction followed by Western blot analysis. NFU4 and NFU5 were enriched in the soluble fraction (Figure 5.3B), similar to the FeS proteins ISU1 and INDH, which have previously been shown to be matrix localised (Wydro et al. 2013). Additionally, GRXS15 was enriched in this fraction. By contrast, the *cox2* subunit of complex IV was as expected enriched in the membrane fraction.

Taken together these data indicate that NFU4 and NFU5, as well as GRXS15, are localised to the matrix of the mitochondria.



**Figure 5.3: NFU4 and NFU5 are localised in the mitochondrial matrix**

**(A)** Western blot analysis of NFU4 and NFU5 in cell fractions. Equal amounts of protein from total leaf extract, chloroplasts and mitochondria were separated by SDS-PAGE and immunolabelled with antibodies against the indicated proteins. NBP35 was used as a cytosolic marker, GRXS15 as a mitochondrial marker and the ponceau stain of Rubisco as a chloroplastic marker.

**(B)** Western blot analysis of NFU4 and NFU5 in soluble and membrane fractions in mitochondria. Purified mitochondria were separated into soluble supernatant and insoluble pellet fractions by freeze-thaw cycles and centrifugation. Protein was then separated by SDS-PAGE and labelled with antibodies against NFU4 and NFU5, GRXS15, ISU1 and INDH (soluble matrix proteins) and cox2 (a membrane bound protein).

**Table 5.2: Summary of mutant alleles**

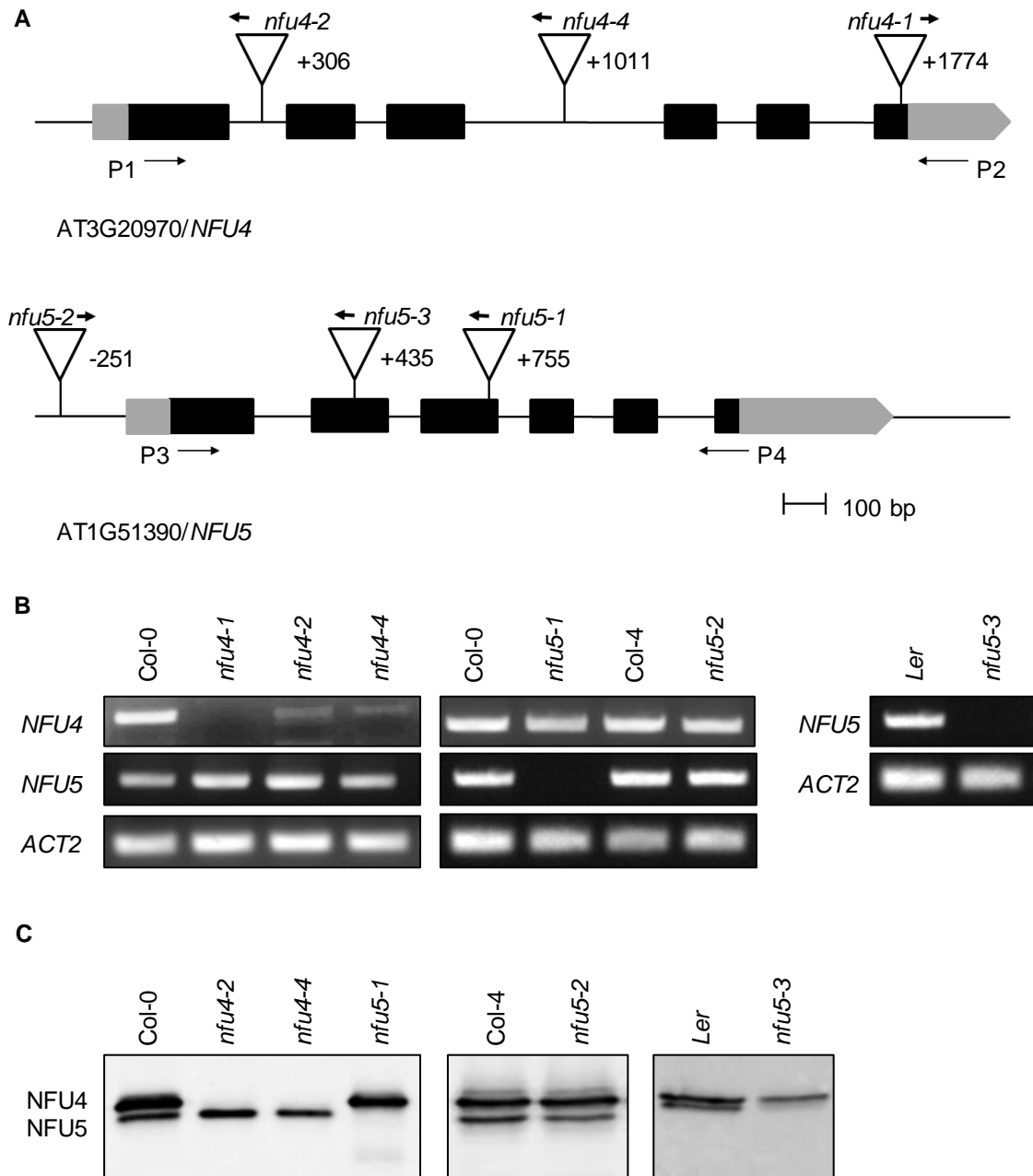
Locus	Allele	Polymorphism	Mutation
AT3G20970 (NFU4)	<i>nfu4-1</i>	SALK_035493	T-DNA insertion (exon 6 +56)
	<i>nfu4-2</i>	SALK_061018	T-DNA insertion (intron 1 +84)
	<i>nfu4-4</i>	SAIL_1233_C08	T-DNA insertion (intron 3 +230)
AT1G51390 (NFU5)	<i>nfu5-1</i>	WiscDsLoxHs069_06B	T-DNA insertion (exon 3 +173)
	<i>nfu5-2</i>	SK24394	T-DNA insertion (promoter -250)
	<i>nfu5-3</i>	GT_3_2834	T-DNA insertion (exon 2 +435)
AT4G22220 (ISU1)	<i>isu1-1</i>	SALK_006332	T-DNA insertion (promoter -64)

### 5.2.3 Genetic analysis of *NFU4* and *NFU5* mutants

To investigate the functions of *NFU4* and *NFU5*, T-DNA insertion mutants were obtained from Arabidopsis stock centres. Homozygous plants were isolated by genotyping PCR analysis and the exact insertion site was determined by sequencing (Figure 5.4A, Table 5.2). Mutants *nfu4-1*, *nfu4-2*, *nfu4-4* and *nfu5-1* are in the Columbia-0 (Col-0) background, *nfu5-2* is in the Columbia-4 (Col-4) background and *nfu5-3* is in the Landsberg (Ler) background. RT-PCR confirmed the absence of full length *NFU4* transcript in *nfu4-1*, and the substantial decrease of full length *NFU4* in *nfu4-2* and *nfu4-4* (Figure 5.4B). The remaining transcript in *nfu4-2* and *nfu4-4* is likely due to contamination of the PCR reaction or spill-over as it is unlikely any product could be amplified over the length of the T-DNA. No *NFU5* transcript was detected in *nfu5-1* and *nfu5-3*. Transcript levels in *nfu5-2* were indistinguishable from WT when looked at by RT-PCR. However, subsequent qRT-PCR (performed by Dr. Robert Green) analysis showed approximately 81% *NFU5* transcript compared to WT. This is consistent with the placement of the T-DNA 251 bp upstream of the ATG start codon in the 3'UTR of the previous gene, *AT1G51380*, which may affect promoter activity of *NFU5*.

Western blot analysis was performed using an antibody that was raised against recombinant *NFU4* but which detects both *NFU4* and *NFU5* (Supplied by Dr. Nicolas Rouhier). In a 15% SDS-polyacrylamide gel both proteins can be easily detected due to their different sizes (Figure 5.4C). Taking into account the predicted mitochondria targeting sequence (MTS) cleavage site (prediction from Uniprot), *NFU4* was calculated to be 25.5 kDa and *NFU5* 23.4 kDa, in agreement with the apparent mobility. Using Western blot analysis, the absence of *NFU4* protein in *nfu4-2* and *nfu4-4* and the absence of *NFU5* protein in *nfu5-1* and *nfu5-3* was confirmed (Figure 5.4C). *NFU5* protein levels in the *nfu5-2* mutant were decreased to around 80% of Col-4. The T-DNA insertion in *nfu4-1* is located in the last exon of *NFU4* and so may result in a truncated protein product. If this truncated protein product is of similar size to *NFU5*, it would be hard to detect by Western blot analysis, and would only be visible in combination with the *nfu5-1* mutant. This double mutant was created, but was not yet analysed, and so *nfu4-1* will no longer be investigated.

These data show that *nfu4-2*, *nfu4-4*, *nfu5-1* and *nfu5-3* can be considered knock-out mutants. *nfu5-2* is a partial mutant. *nfu4-1* contains no full-length transcript but may retain truncated protein product.

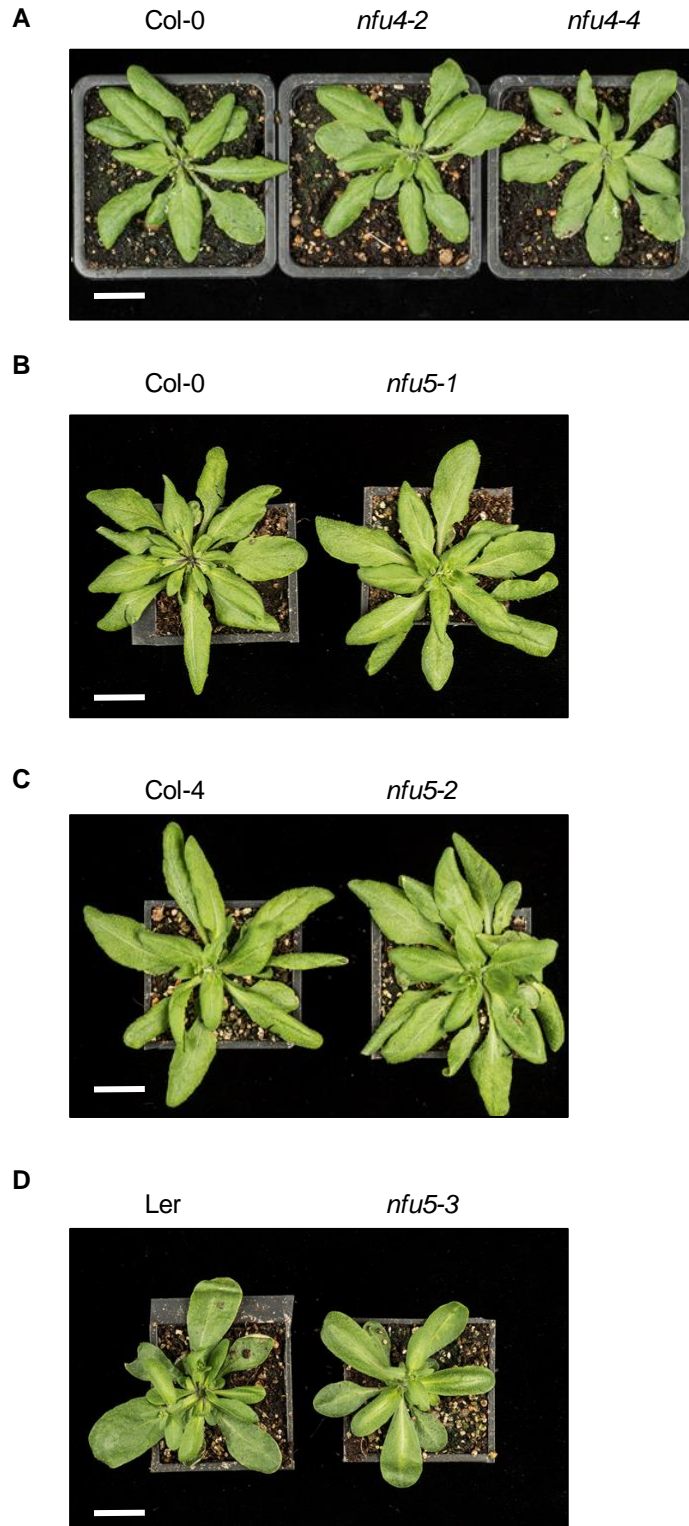


**Figure 5.4: Genetic analysis of *Arabidopsis thaliana* mutants in *NFU4* and *NFU5***

**(A)** Gene model of *NFU4* and *NFU5* and the positions of T-DNA insertions. Black bars represent exons, grey bars are the 5' and 3' untranslated regions of the transcript. Triangles represent T-DNA insertions, their orientation is marked with an arrow to indicate the outward facing left border primer. Primers P1 – P4, used for RT-PCR, are also marked by arrows.

**(B)** Transcript levels of *NFU4* and *NFU5* in wild-type seedlings and T-DNA mutants. RNA was isolated from 2-week-old seedlings, reverse transcribed to cDNA with an oligo dT primer and PCR was performed with primer pairs P1 and P2 for *NFU4*, and P3 and P4 for *NFU5* (see A). *ACTIN2* was used as a control for the amount of cDNA template.

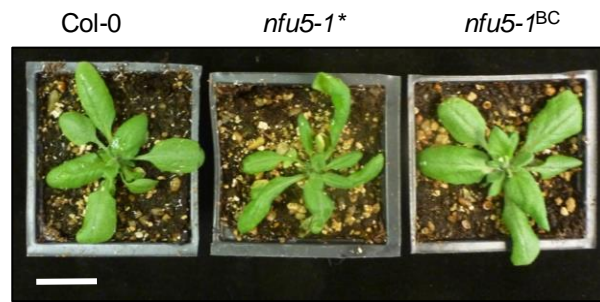
**(C)** *NFU4* and *NFU5* protein levels shown by SDS-PAGE followed by immunolabelling of *NFU4* and *NFU5* proteins in purified mitochondria samples isolated from 2-week-old seedlings (left and middle panel) and in total protein from 7-week-old plants (right-hand panel).



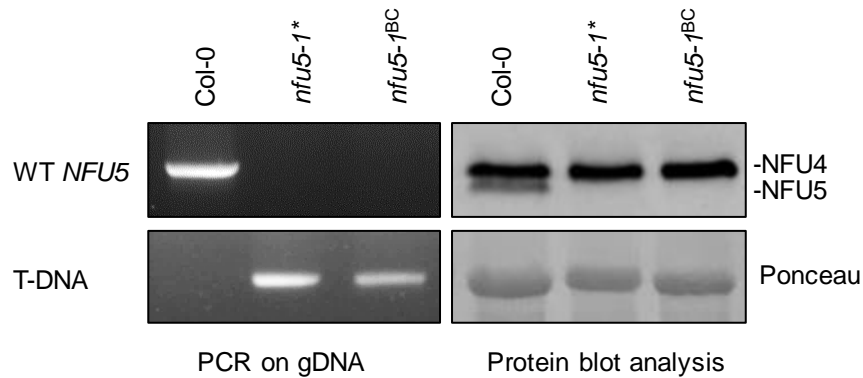
**Figure 5.5: Growth analysis *NFU4* and *NFU5* T-DNA insertion lines**

(A-D) Growth phenotype of 4-week-old plants of T-DNA insertion lines and corresponding wild types. *nfu4-1*, *nfu4-2*, *nfu4-4* and *nfu5-1* are in the Columbia 0 (Col-0) background, *nfu5-2* is in the Columbia 4 (Col-4) background and *nfu5-3* is in the Landsberg (Ler) background. Scale bar = 2 cm.

**A**



**B**



**Figure 5.6: Isolation of *nfu5-1* line**

**(A)** Growth phenotype of 5-week-old plants of Col-0, *nfu5-1\** (original T-DNA insertion line) and *nfu5-1<sup>BC</sup>* (re-isolated *nfu5-1*). Scale bar = 2 cm.

**(B)** Genotype and protein levels from lines in (A). The presence of the T-DNA insertion was shown by genotyping PCR using primers from the left border of the T-DNA and a genomic primer, and the absence of WT *NFU5* by two genomic primers. *NFU4* and *NFU5* protein levels shown by SDS-PAGE followed by immunolabelling of *NFU4* and *NFU5* proteins from total leaf extracts.

#### 5.2.4 *NFU4* and *NFU5* mutants display no growth phenotype suggesting functional redundancy

In order to evaluate the importance of *NFU4* and *NFU5* for plant growth, all T-DNA insertion lines were grown on soil under long-day conditions (Figure 5.5A-D). The *nfu4-2* and *nfu4-4* insertion mutants had no apparent growth phenotype. The *nfu5-2* and *nfu5-3* insertion mutants appear indistinguishable from their respective wild types.

Initially *nfu5-1* appeared to have a growth phenotype (*nfu5-1\**). The plants had narrow leaves and were smaller than WT (Figure 5.6A). However, during subsequent experiments where this insertion mutant was crossed with different lines I noticed that the growth phenotype did not segregate with the *nfu5-1* homozygous genotype. This suggested that the growth phenotype was due to a secondary mutation, unrelated to the *NFU5* gene. *nfu5-1* homozygous plants re-isolated from a back cross (*nfu5-1<sup>BC</sup>*) were indistinguishable from WT (Figure 5.6A, B). This showed that the *nfu5-1* insertion mutant has no growth phenotype, in agreement with normal growth in the *nfu5-3* mutant.

These data suggest that neither NFU4 nor NFU5 individually have a critical role under normal growth conditions.

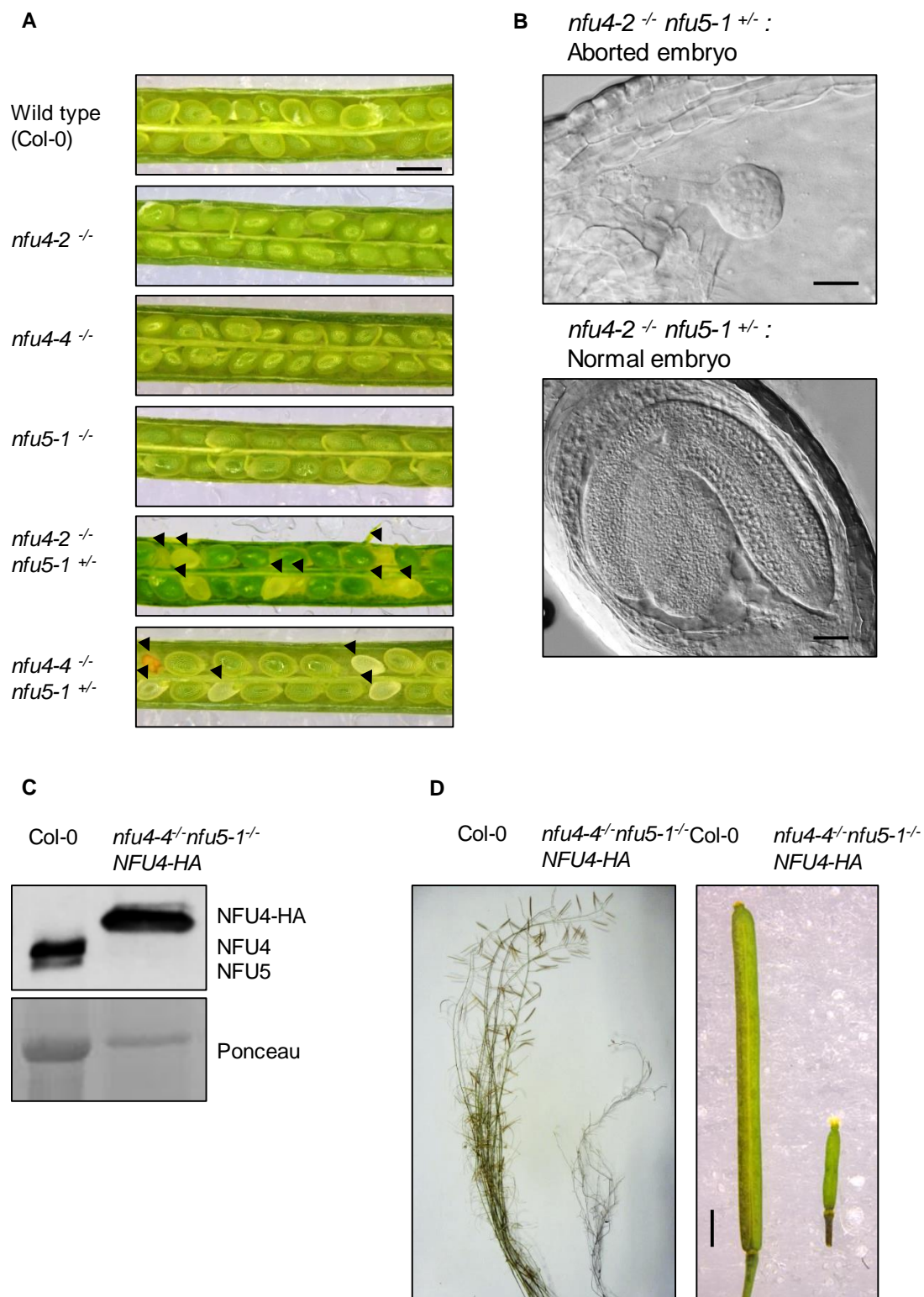
#### 5.2.5 *NFU4* and *NFU5* are essential for embryo development.

Given the lack of a clear growth phenotype in the single mutants, it is likely that NFU4 and NFU5 play redundant functions. Therefore, efforts were made to create a double mutant between knock-out alleles, by reciprocal crosses between *nfu5-1* and either *nfu4-2* or *nfu4-4*. However, in the F2 generation no double mutants could be isolated. Therefore, I looked in the siliques of *nfu4-2<sup>-/-</sup>nfu5-1<sup>+/-</sup>* and *nfu4-4<sup>-/-</sup>nfu5-1<sup>+/-</sup>* plants to see if embryo development was affected. Approximately one quarter of the immature seeds were white, consistent with an expected ratio of 3:1, indicating embryo abortion for *nfu4-2<sup>-/-</sup>nfu5-1<sup>-/-</sup>* and *nfu4-4<sup>-/-</sup>nfu5-1<sup>-/-</sup>* individuals (Figure 5.7A, Table 5.3). Col-0 and single mutants involved in the crosses (*nfu4-2*, *nfu4-4* and *nfu5-1*), differed significantly from a 3:1 ratio, and so show no evidence of significant embryo abortion. Therefore, it is likely that a complete lack of the NFU4 and the NFU5 protein is lethal and results in embryo abortion.

In order to further investigate the role of NFU proteins in embryo development embryos were cleared using Hoyer's solution and imaged Normanski differential interference contrast (DIC) microscopy. To find the terminal phenotype of the aborted embryos I looked at both green (normal) and white (aborted) ovules, from the same silique, at a



Chapter 5 Functional characterisation of FeS assembly proteins in *Arabidopsis thaliana* mitochondria



**Figure 5.7: Investigation of *NFU4 NFU5* double mutant**

**(A)** Silique pictures, showing seedling development, in wild type (Col-0), *nfu4* and *nfu5* single mutants as well as *nfu4-2<sup>-/-</sup> nfu5-1<sup>+/-</sup>* and *nfu4-4<sup>-/-</sup> nfu5-1<sup>+/-</sup>*. Scale bar = 500 µm. Black arrows indicate aborting seeds.

**(B)** An aborted and healthy embryo from the silique of *nfu4-2<sup>-/-</sup> nfu5-1<sup>+/-</sup>*. Siliques were cleared with Hoyer's solution and imaged with DIC microscopy. Scale bar = 25 µm (top panel) and 50 µm (bottom panel). Performed with Dr. Vasilios Andriotus.

**(C)** Western blot analysis of *nfu4-4<sup>-/-</sup> nfu5-1<sup>-/-</sup> NFU4-HA*, showing levels of NFU4, NFU5 and NFU4-HA protein.

**(D)** Growth phenotype and silique morphology of Col-0 and *nfu4-4<sup>-/-</sup> nfu5-1<sup>-/-</sup> NFU4-HA*. Scale bar = 1 mm.

**Table 5.3: Embryo abortion counts of *NFU* mutants**

$\chi^2$ -value for embryo abortion calculated on an expected ratio of 3:1.

	Normal (%)	Aborted (%)	Not fertilised (%)	n	$\chi^2$ -value
Wild type (Col-0)	99.0	0.7	0.3	408	128 (p<0.001)
<i>nfu4-2<sup>-/-</sup></i>	97.8	2.2	0.0	359	99 (p<0.001)
<i>nfu4-4<sup>-/-</sup></i>	99.2	0.5	0.3	386	123 (p<0.001)
<i>nfu5-1<sup>-/-</sup></i>	99.5	0.0	0.5	430	142 (p<0.001)
<i>nfu4-2<sup>-/-</sup> nfu5-1<sup>+/-</sup></i>	76.8	23.2	0.0	384	0.7
<i>nfu4-4<sup>-/-</sup> nfu5-1<sup>+/-</sup></i>	72.8	26.1	1.1	448	0.5

late stage of development in *nfu4-2<sup>-/-</sup>nfu5-1<sup>+/-</sup>* (Performed with help from Dr. Vasilios Andriotus). While the green ovules contained embryos at the mature green stage, the white ovules contained an embryo arrested at the globular stage (Figure 5.7B). This suggests that the globular phase of development is the terminal phenotype. Plants without *NFU4* and *NFU5* cannot survive beyond this point.

In order to confirm that the embryo lethality phenotype is caused by a lack of functional NFU protein, the double mutant was complemented with an C-terminally HA tagged *NFU4* construct (pBIB plasmid, 35S promoter, from Dr. Nicolas Rouhier). *nfu4-4* + *NFU4-HA* individuals were crossed with *nfu5-1*. One complemented individual, *nfu4-4<sup>-/-</sup>nfu5-1<sup>-/-</sup>NFU4-HA*, out of ~40 individuals, was isolated in the F2 generation (Figure 5.7C). The individual was small and stunted and had abnormally small silique. In addition, the silique had not been fertilised indicating it is male sterile (Figure 5.7D). The individual's ability to survive and reach reproductive age demonstrates partial complementation of the embryo lethal phenotype and subsequent vegetative growth. However, *NFU4-HA* cannot replace the function of endogenous *NFU4* during reproduction.

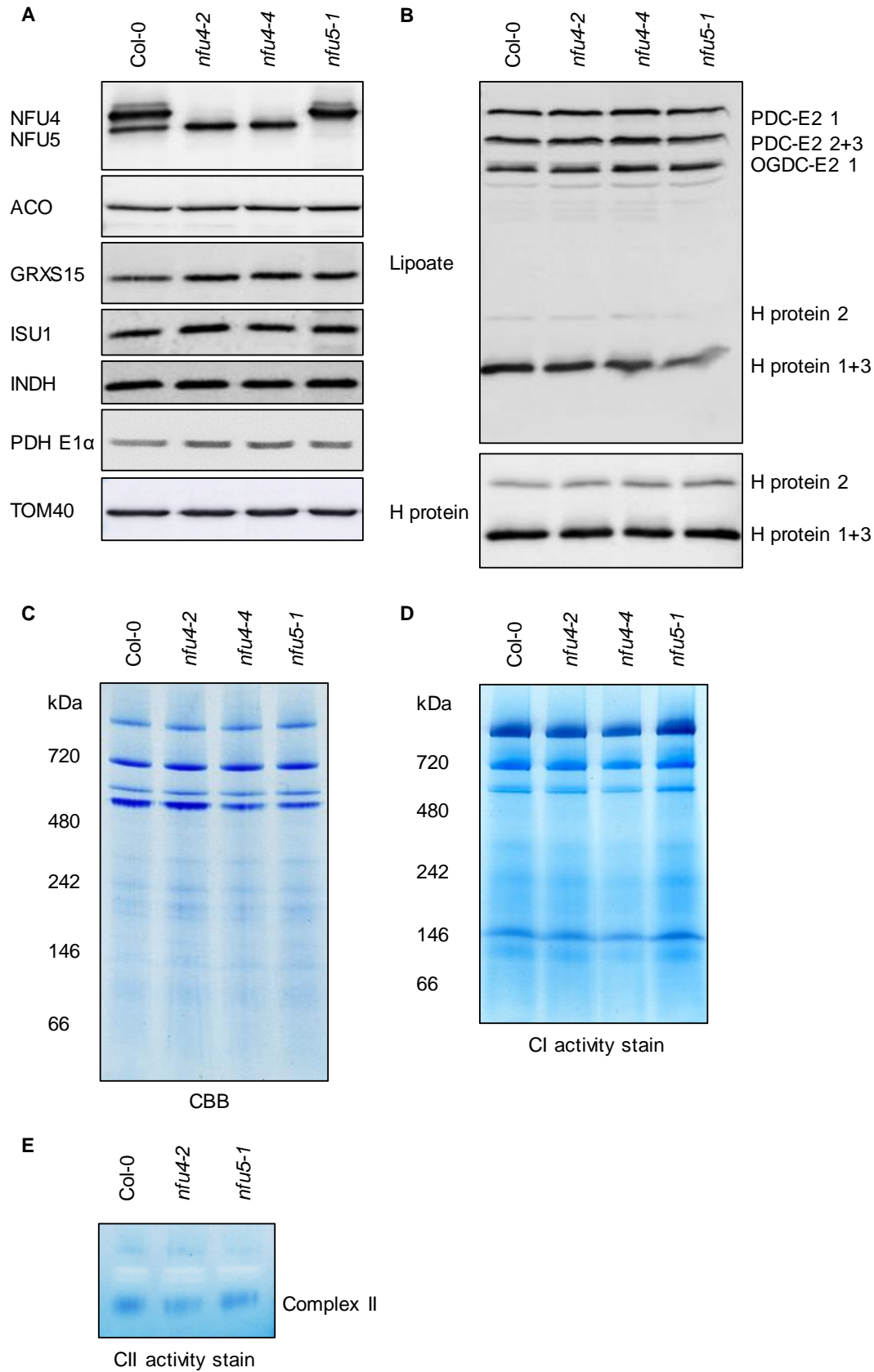
Overall, these data suggest a critical function for the NFU proteins during a particular growth stage.

### 5.2.6 Single mutants of *NFU4* and *NFU5* have no noticeable FeS assembly phenotype

In order to characterise the role of *NFU4* and *NFU5* in FeS assembly, FeS containing proteins and enzymes were investigated. To study the effect of *NFU4* and *NFU5* mutants on a range of FeS containing proteins, Western blot analysis was performed on isolated mitochondria. No decrease in protein levels of the FeS binding proteins aconitase, ISU1, INDH and GRXS15, nor on the E1 $\alpha$  subunit of pyruvate dehydrogenase was seen (Figure 5.8A).

Lipoate cofactor synthesis is dependent of FeS assembly, because lipoate synthase requires an FeS cluster. Therefore, mutants in FeS assembly can have lipoate defect phenotypes. An effect on lipoate synthesis is one of the main defects seen in human and *Saccharomyces NFU1* mutants so it was decided to study lipoate containing proteins in more detail. In *Arabidopsis*, there are seven lipoate containing proteins: three isoforms of the H-protein of the glycine cleavage complex (H1, At2g35370; H2, At2g35120 and H3, At1g32470), three E2 proteins from the PDC complex (E2-1, At3g52200; E2-2, At3g13930 and E2-3, At1g54220) and an E2 protein from the OGDH enzyme complex (At5g55070) (Stroeher et al. 2015). Using antibodies specific to lipoic acid and to H-

Chapter 5 Functional characterisation of FeS assembly proteins in *Arabidopsis thaliana* mitochondria



**Figure 5.8: Western blot and BN-PAGE analysis of *NFU4* and *NFU5* single mutants**

**(A)** Levels of the mitochondrial proteins NFU4, NFU5, aconitase (ACO), glutaredoxin S15 (GRXS15), FeS cluster assembly protein 1 (ISU1), FeS protein required for NADH Dehydrogenase homolog (INDH), E1 $\alpha$  subunit of pyruvate dehydrogenase (PDH) and translocase of the outer membrane 40 (TOM40) in wild type (Col-0), *nfu4* and *nfu5* single mutants, as indicated. Mitochondria were purified from 2-week-old seedlings and analysed by immunoblotting with specific antibodies.

**(B)** Lipoate-binding proteins in mitochondria of wild type (Col-0) and *nfu4* or *nfu5* single mutants. Purified mitochondria from 2-week-old seedlings were subjected to immunolabeling with antibodies against lipoate and H-protein.

**(C)** Respiratory complexes I, III and V visualised by BN-PAGE and Coomassie staining in wild type (Col-0), *nfu4* and *nfu5* single mutants, as indicated. Mitochondria were purified from 2-week-old seedlings.

**(D)** Respiratory complex I separated by BN-PAGE followed by activity staining with NADH and NBT of wild type (Col-0), *nfu4* and *nfu5* single mutants, as indicated. Mitochondria were purified from callus culture.

**(E)** Respiratory complex II separated by BN-PAGE followed by activity staining with PMS, Succinate and NBT of wild type (Col-0), *nfu4* and *nfu5* single mutants, as indicated. Mitochondria were purified from callus culture.

proteins, no decrease in any these proteins or cofactors (Figure 5.8B), except for a slight decrease in the lipoate bound to the H2 protein in *nfu5-1*.

The abundance of respiratory complexes in *nfu4-2*, *nfu4-4* and *nfu5-1* single mutants was analysed using BN-PAGE and Coomassie staining (Figure 5.8C). No decreases of any respiratory complexes were observed. Activity staining with NBT/NADH, to visualise complex I levels, again showed no decreases (Figure 5.8D). Complex II levels were visualised with an in gel histochemical stain. Again, no decrease in the *nfu4-2* or *nfu5-1* mutants was seen (Figure 5.8E).

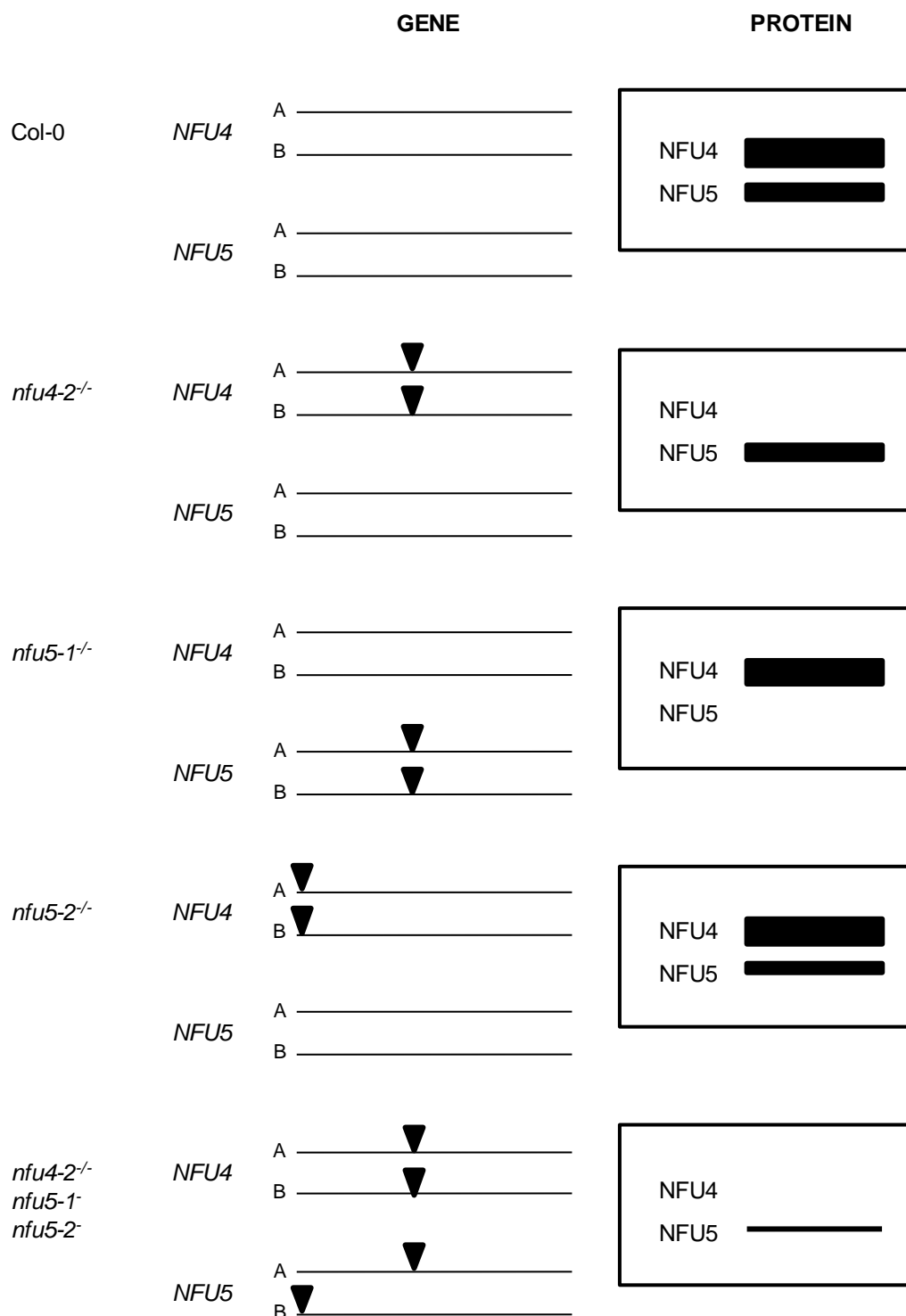
Taken together these data suggests that neither *NFU4* nor *NFU5* single mutants have observable defects in FeS assembly. FeS containing enzymes affected in human and *Saccharomyces NFU1* mutants were not decreased in activity in *Arabidopsis thaliana NFU4* or *NFU5* single mutants.

### 5.2.7 Hemizygous *NFU* mutant has a decrease in H-protein

Single mutants in *NFU4* and *NFU5* do not have observable growth or FeS cluster assembly defects. Double knock-out mutants of *NFU4* and *NFU5* are embryo lethal, making further study of their function problematic. Therefore, in order to study the role of *NFU4* and *NFU5*, a hemizygous mutant was created. This was done by combining in one individual three T-DNA insertions: *nfu4-2*, *nfu5-1* and *nfu5-2* (Figure 5.9). This line has no *NFU4* protein and low levels of *NFU5* protein (Figure 5.10B). BN-PAGE and Western blot analysis were performed to assess the levels of respiratory complexes and FeS proteins. Additionally, a *nfu4-2-nfu5-2* double mutant, with no *NFU4* protein and ~80% *NFU5* protein, was analysed.

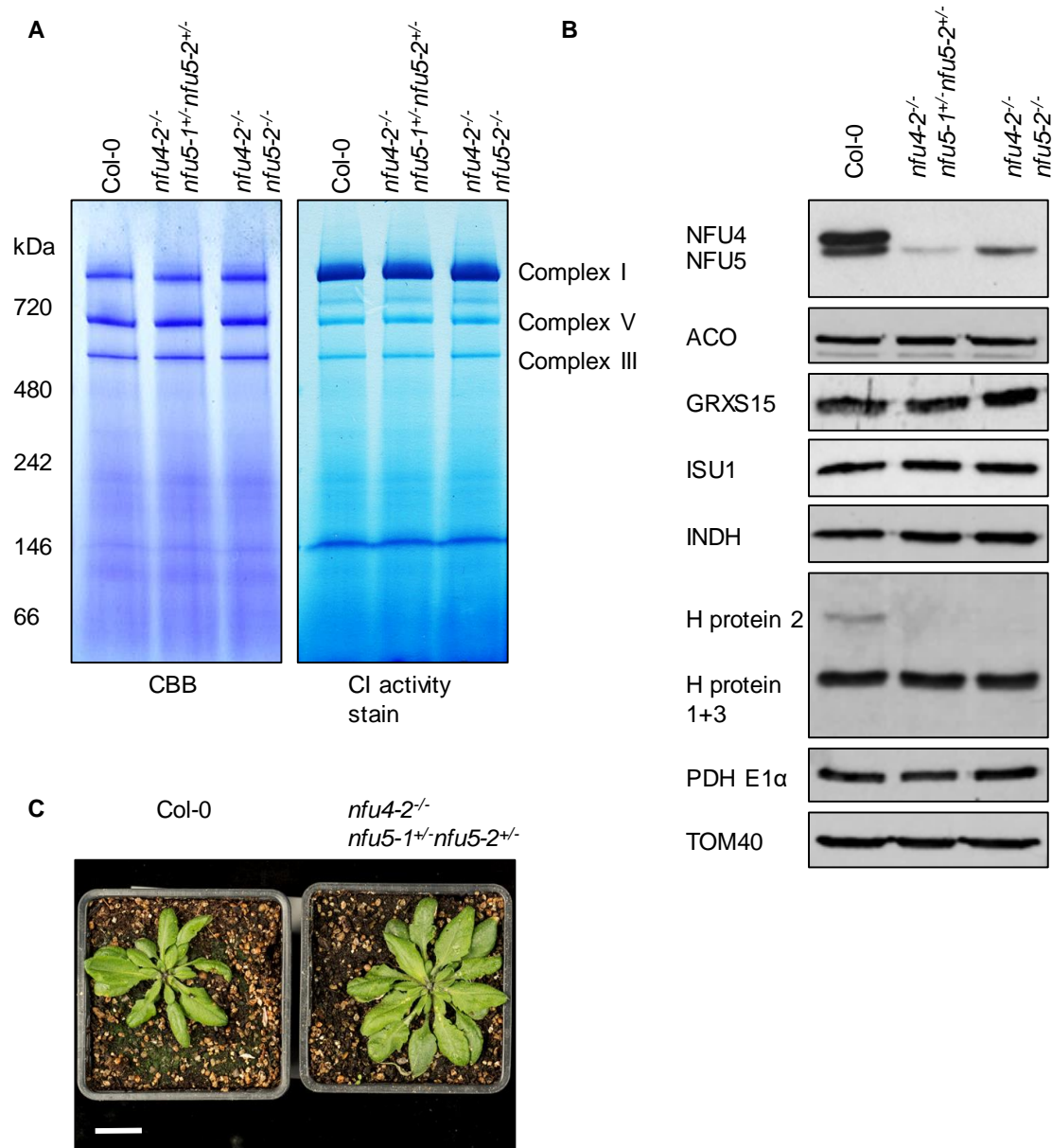
Similarly to what was seen in the single mutants, no defect was seen in the levels of respiratory complexes (Figure 5.10A). The levels of aconitase, ISU1, INDH and GRXS15, and the E1 $\alpha$  subunit of pyruvate dehydrogenase were unaffected (Figure 5.10B). However, a substantial decrease was seen in the levels of the H2 protein of GDC in both the hemizygous and *nfu4-2-nfu5-2* mutants.

Despite having very low levels of NFU protein growth was unaffected compared to WT (Col-0) (Figure 5.10C).



**Figure 5.9: Creation of the hemizygous *NFU* mutant**

Schematic detailing genetic composition of the hemizygous mutant. Black triangles indicate T-DNA insertion. The two black lines under the heading protein depict the relative levels of NFU4 and NFU5 proteins.



**Figure 5.10: BN-PAGE and Western blot analysis of hemizygous and *nfu4-2 nfu5-2* double mutants**

**(A)** Respiratory complexes I, III and V separated by BN-PAGE and visualised with Coomassie staining (left-hand panel) and by complex I NADH/NBT activity staining (right-hand panel) in wild type (Col-0), hemizygous and *nfu4-2 nfu5-2* double mutant. Mitochondria were purified from callus culture.

**(B)** Levels of the mitochondrial proteins NFU4, NFU5, ACO, GRXS15, ISU1, INDH, H-protein, E1α subunit of pyruvate dehydrogenase (PDH) and TOM40 in wild type (Col-0), hemizygous and *nfu4-2 nfu5-2* double mutant as indicated. Mitochondria were purified from callus culture.

**(C)** Growth phenotype of 4-week-old plants of Col-0 and hemizygous mutant. Scale bar = 2 cm



### 5.2.8 *GRXS15* mutants do not have a strong FeS assembly phenotype

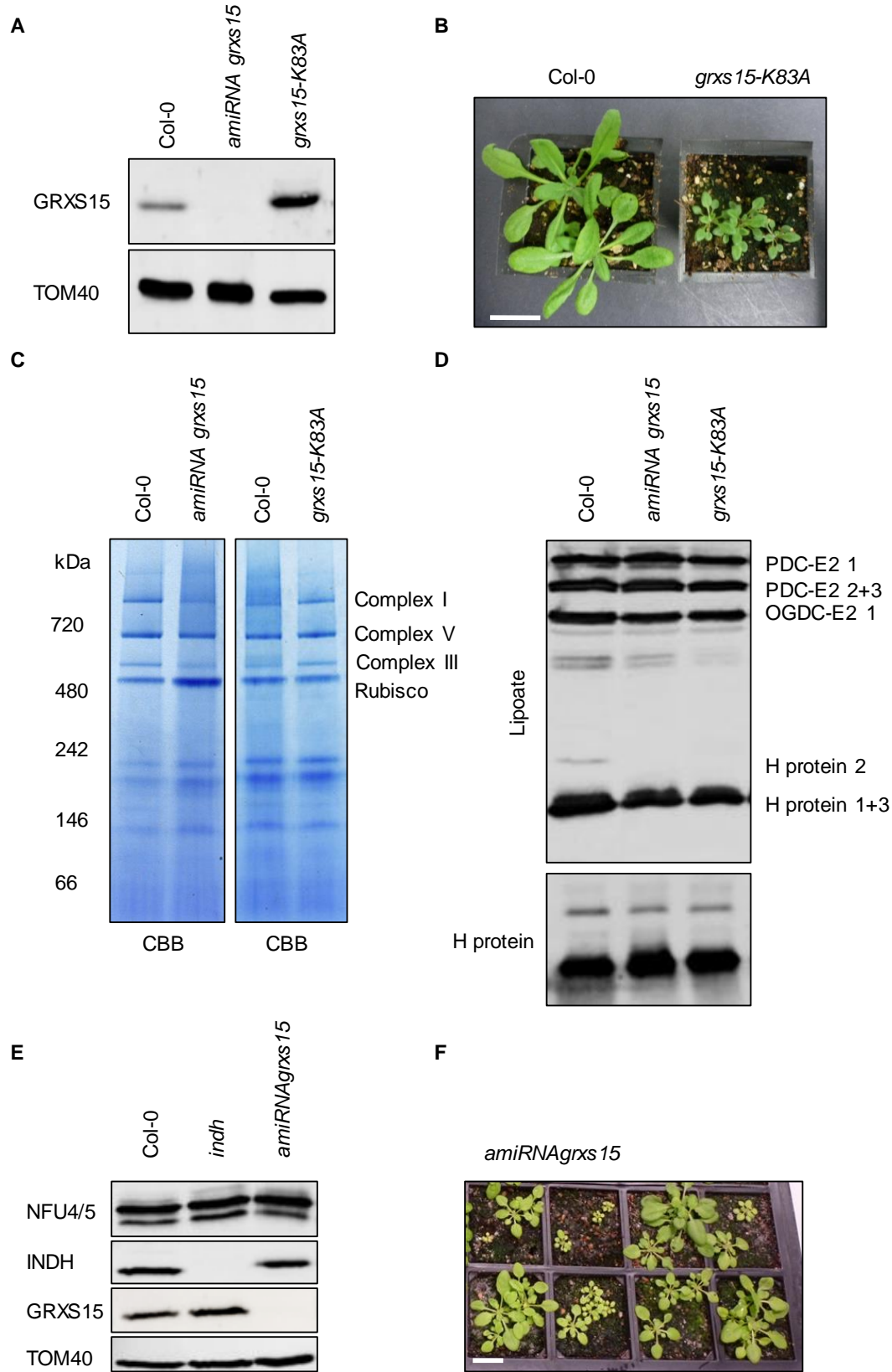
Recent studies on *GRXS15* in *Arabidopsis thaliana* conclude that it is localized in mitochondria and plays a role in FeS assembly in the mitochondria (Moseler et al. 2015; Stroehler et al. 2015). However, the two reports described different mutant lines, and the findings differ, for example with respect to aconitase activity. To obtain clarity about the role of *GRXS15* in FeS cluster assembly in plants, the two *GRXS15* mutants were grown side-by-side and further characterised. This work was initially undertaken with Dr. Inga Kruse and Dr. Janneke Balk. Some of the ideas and conclusions here were discussed in the thesis of Dr. Inga Kruse, as was the lipoate blot in figure 5.11E, which was undertaken jointly.

First, the RNAi knockdown line described in Stroher et al 2015, *amiRNA grxs15*, was obtained from the group of Prof. Harvey Millar. This line has been shown by multiple reaction monitoring analysis to contain approximately 3% of *GRXS15* compared to WT (Stroehler et al. 2015). Consistent with this, Western blot analysis does not detect any *GRXS15* protein in this line, as it is below the detection limit (Figure 5.11A, Stroehler et al. 2015). The growth phenotype of *amiRNA grxs15* is variable, as often seen in RNAi lines, but smaller than WT (Figure 5.11F, Dr. Inga Kruse thesis, Stroehler et al. 2015).

The second line was obtained from the group of Prof. Andreas Meyer. The line was created by complementing a knock out line, *grxs15-3*, with a construct containing *GRXS15* with a key catalytic residue, lysine 83, mutated to alanine, giving *grxs15-3 UBQ10<sub>pro</sub>: GRXS15 K<sub>83</sub>/A*. Lys83 is involved in glutathione binding. The mutant line will henceforth be referred to as *grxs15-K83A*. *GRXS15* protein levels are higher than in wild type because the mutant gene is expressed using the *UBIQUITIN* promoter (Figure 5.11A and Moseler et al. 2015). The growth phenotype of *grxs15-K83A* is small and chlorotic (Figure 5.11B, Moseler et al. 2015).

The abundance of respiratory complexes in both *GRXS15* mutants was analysed using BN-PAGE and Coomassie staining (Figure 5.11C). No decreases of any respiratory complexes were observed in either mutant.

Previously published results suggested lipoate containing proteins could be affected in the *amiRNA grxs15* mutant (Stroehler et al. 2015). In order to confirm this result, and test to see if the same was true in *grxs15-K83A*, Western blot analysis was performed with antibodies against lipoate and H-protein. No decrease was seen in the levels of H-protein, unlike in Stroher et al. 2015, however a decrease was seen in the lipoate bound



**Figure 5.11: Analysis of *grxs15-K83A* and *amiRNA grxs15* mutants**

**(A)** Protein levels of GRXS15 in Col-0, *grxs15-K83A*, and *amiRNAgrxs15*. Mitochondria were purified from 4-week-old plants and analysed by immunoblotting with specific antibodies. TOM40 was used as a marker protein for equal transfer and loading.

**(B)** Growth phenotype of 3-week-old *grxs15-K83A* mutant and corresponding wild type, Col-0. Scale bar = 2 cm.

**(C)** Respiratory complexes I, III and V separated by BN-PAGE and visualised with Coomassie staining in Col-0, *grxs15-K83A* and *amiRNAgrxs15*. Mitochondria were purified from 4-week-old plants.

**(D)** Lipoate and H-protein levels on indicated proteins in Col-0, *grxs15-K83A*, and *amiRNAgrxs15*. Mitochondria were purified from 4-week-old plants and analysed by immunoblotting

**(E)** Levels of the mitochondrial proteins NFU4, NFU5, GRXS15, INDH and TOM40 in Col-0, *indh* and *amiRNAgrxs15*. Mitochondria were purified from either callus (lane 1 and 2) or 4-week-old plants (lane 3) and analysed by immunoblotting with specific antibodies.

**(F)** Growth phenotype of 4-week-old *amiRNAgrxs15*. Scale bar = 2 cm.

H-protein 2 of GDC for both mutants (Figure 5.11E). Other lipoate containing proteins were unaffected. These data indicate that the lipoate bound to H-protein 2 is decreased in both *grxs15* mutants, and confirms a defect in lipoate metabolism in *GRXS15* mutants, even if the pattern of defect is slightly different from that in Stroher et al. 2015.

Western blot analysis was performed to investigate if the absence of *GRXS15* affected other FeS carrier proteins. The *amiRNA grxs15* line showed normal levels of FeS binding proteins NFU4, NFU5 or INDH (Figure 5.11E), suggesting that these act independently. The lack of effect on NFU4 protein levels is unexpected given that Stroher et al. 2015 found that in *GRXS15* knock out lines NFU4 protein is decreased 1.7-fold (Stroehrer et al. 2015).

Taken together these data show *grxs15* mutants have a strong growth phenotype but a mild FeS assembly defect. An FeS defect was unobservable except from the decrease in lipoate bound to the H-protein 2 of GDC.

### 5.3 Discussion

Here, I found that double knock out mutants of *NFU4* and *NFU5* were embryo lethal. This shows that, together, they play an essential role in FeS assembly. However, it does make assigning their exact role in the process problematic. This continues a common pattern in FeS assembly research. FeS assembly proteins are highly conserved across the eukaryotes and are often essential, therefore mutants are often non-viable. This fact has frustrated efforts to study FeS proteins in many different model systems. Innovative solutions have been needed to get around this problem and study protein function. In *Arabidopsis thaliana*, there are numerous mutants in FeS assembly that have been found to be, at least partially, embryo lethal. These include, from the mitochondrial ISC pathway, NFS1 (Frazzon et al. 2007); IBA57 (Waller et al. 2012); *GRXS15* (Moseler et al. 2015); Frataxin (Busi et al. 2006; Vazzola et al. 2007), ISU1 (Janneke Balk, unpublished data); NFU4 and NFU5 (this chapter), the CIA pathway NBP35 (Bych et al. 2008b; Kohbushi et al. 2009); DRE2 (Bernard et al. 2013); (Varadarajan et al. 2010); AE7 (Luo et al. 2012), NAR1 (Cavazza et al. 2008; Luo et al. 2012) and CIA1 (Luo et al. 2012), and the chloroplastic SUF pathway SUFE1 (Xu & Møller 2006); SUFE3 (Narayana Murthy et al. 2007); SUFC/NAP7 (Xu & Møller 2004) and SUFD/NAP7 (Hjorth et al. 2005). This demonstrates that my findings with *NFU4* and *NFU5* fit well with previous observations of FeS assembly in plants.

Efforts to overcome embryo lethality phenotypes to study protein function centre around two main strategies. The first is to decrease the amount of protein present, either through

RNAi or different insertion mutants, to low enough levels that a phenotype is seen, but not so much that the embryo is unviable. This is the strategy I decided on in creating the *NFU* hemizygous mutant. Western blot analysis showed a dramatic decrease in total *NFU* protein levels (Figure 5.10B). However, the phenotypic effects were mild: no growth defect was observed and most FeS proteins were unaffected. Some decrease was seen in lipoate containing H2 protein of the GDC. Further work will focus on performing enzyme assays on these lines to see if complex II, PDH or GDC are affected. A similar result is seen in investigating *isu1-1*. Protein levels could be decreased to approximately 18% without any observable phenotype (Figure 5.2B). This has been observed before for FeS proteins. For example, in studying NBP35, RNAi could decrease NBP35 protein to approximately 30% of WT, without any detrimental effects on growth or FeS enzymes (Bastow et al. 2017). Clearly, in the study of FeS proteins, this strategy has limited success, as *Arabidopsis thaliana* can survive with very low levels of these essential proteins. This raises the puzzling question of why the plant produces high levels of them when low levels suffice.

Due to the modular nature of the NifU protein in *Azotobacter vinelandii*, with homology to both *Isu1* and *Nfu* proteins (Agar et al. 2000) double mutants were made in *Saccharomyces cerevisiae*. These double mutants had more severe defects than seen in the single mutants (Schilke et al. 1999). Combining *isu1* and *nfu4* or *nfu5*, mutants in *Arabidopsis thaliana* may have a similar effect. Despite individual mutants having enough protein so that no phenotype is seen, it is hoped that in the double mutant, with decreased protein of both *ISU1* and *NFU4* or *NFU5*, would allow a mutant phenotype to be seen. This work was started, but was not completed in time for inclusion in this thesis.

The second strategy involves complementing a knock out mutant with a version of the protein with a key amino acid residue substituted. This allows minimum functionality required to avoid lethality, but also a sufficient decrease in protein function leading to a phenotype that is possible to study. This strategy was effectively adopted in studying NBP35 where knock out mutants of *NBP35* were complemented with a copy of *NBP35* with an amino acid substitution at cysteine 14 (Bastow et al. 2017). This allowed phenotypic characterisation of the mutant. This strategy was also used, as previously mentioned in studying *GRXS15* with the *grxs15-K83A* mutant (Moseler et al. 2015). Time constraints prevented the use of this strategy in studying *NFU4* and *NFU5*, but should certainly be considered for future work.

Before this work was undertaken the role of *GRXS15* in plants was uncertain. It was suggested to have a role in FeS assembly in the mitochondria based on FeS defects

observed in mutants in homologs, for example in *Saccharomyces cerevisiae* Grx5 (Rodríguez-Manzanque et al. 2002) and in *Danio rerio* Grx5 (Wingert et al. 2005). AtGRXS15 could complement yeast *grx5* $\Delta$  (Cheng 2008; Moseler et al. 2015) suggesting a similar role. Additionally, numerous Grx proteins have experimentally been shown to coordinate FeS clusters, starting with the human dithiol Grx2 (Lillig et al. 2005). Recent papers suggest a clear role in FeS assembly in *Arabidopsis thaliana* (Stroeher et al. 2015; Moseler et al. 2015). This conclusion is based upon strong *in vitro* but unconvincing *in planta* evidence. My results suggest GRXS15 does not play a central role in the mitochondrial ISC pathway. In two severe mutants, no defects could be found in FeS containing proteins. There remains a clear mismatch in the severity of the growth phenotype in *GRXS15* mutants, and the mildness of the FeS defect observed. The implications of this for the model of the mitochondrial ISC pathway will be discussed in the general discussion.

In this section of work, I aimed to study the late acting FeS carrier proteins in mitochondria, in order to clarify their role and better understand FeS provision to the respiratory complexes. My results suggest that GRXS15 does not play as central a role as is currently thought, and that the NFU4 and NFU5 proteins are essential.

## 6 Discussion

### 6.1 The function of Ind1 in complex I assembly

#### 6.1.1 The role of Ind1 is evolutionarily conserved but the molecular function is unclear

The function of Ind1 was first investigated in the aerobic yeast *Yarrowia lipolytica* in 2008 (Bych et al. 2008). Due to its similarity to cytosolic FeS assembly scaffolds NBP35 and CFD1, and its ability to coordinate a [4Fe-4S] cluster *in vitro* it was proposed to insert FeS clusters into complex I, but not other FeS proteins. Along with the human homolog NUBPL, it has been consistently labelled as the provider of FeS clusters in complex I ever since. However, there are three problems with this view that need addressing. The first is that in both *Yarrowia lipolytica*, human and *Arabidopsis thaliana*, when Ind1/NUBPL/INDH is knocked out, residual complex I activity remains. For example, in *Yarrowia lipolytica*, between 20-30% of complex I persists in *ind1*Δ. Enzymatic activity of complex I, measured using an assay which utilises the entire FeS chain, matches up with complex I levels. This means the complex I that is present is fully functional. In *Arabidopsis thaliana*, in callus tissue, complex I is clearly visible in the *indh* mutant (chapter 4). These data strongly suggest that other factors also contribute to supply FeS clusters. Secondly, Ind1 has never been found in direct association with a complex I assembly intermediate or subunit (for example using the complexome studies carried out in humans: (Guerrero-Castillo et al. 2016); or in *Arabidopsis thaliana*: (Senkler et al. 2016). It would be assumed that to deliver FeS clusters it would have to physically associate with complex I subunits at some point in the assembly process. However, it may simply be that this interaction is transient and cannot be detected using the techniques commonly employed. Finally, recent studies have implicated other FeS assembly proteins, namely HSC20, as the only provider of FeS clusters to complex I (Maio et al. 2017) (see section 6.2.2). In summary, Ind1 is unlikely to be the sole provider of FeS clusters to complex I, and no direct experimental evidence has been found that FeS clusters are inserted into complex I by Ind1.

It is therefore timely to ask the question: what is the molecular function of Ind1? In an attempt to answer this I will take into account the published data across many species, and the data presented in this thesis. The assumption is that Ind1 has the same mode of action in every eukaryotic species. This seems reasonable given that in all three species in which it has been characterised there is a specific and substantial complex I

defect (Bych et al. 2008; Sheftel et al. 2009; Wydro et al. 2013). However, there may be some differences seen in plants (see section 6.1.3).

### 6.1.2 A direct role in complex I assembly

The first question is whether Ind1 plays a direct or indirect role in complex I assembly. If Ind1 plays a direct role in complex I assembly, either inserting FeS cluster or some other role, the natural question is at which stage in the assembly process. This can be inferred from assembly intermediates. Note, the exact mode of action may differ between yeast and humans, as they have different assembly pathways (Vogel et al. 2007).

In *Yarrowia lipolytica*, *ind1Δ*, as well as *ind1Δ* variants, (chapter 3), accumulate a Q module assembly intermediate, but can still make a larger assembly intermediate, as well as a small amount of mature complex I. This could imply a defect in the Q module, whereby the subunits are attached but cannot have their cluster inserted, although a subset do, and make the larger assembly intermediate and mature complex I. Alternatively, it could indicate a defect in the N module, where Q module is made but is in excess as there is not enough N module to attach to. It is also possible there is a membrane arm defect, and so matrix arm assembly intermediates are in excess. Distinguishing between these possibilities can be helped by comparing the defects seen in *ind1Δ* with other examples in the literature. For example, the pattern of assembly intermediates is reminiscent of those seen in complexome sequencing of *n7bmΔ* (human NDUFA12), *n7bmlΔ* (human assembly factor NDUFAF2), and to a lesser extent *nummΔ* (human NDUFS8) (Kmita et al. 2015 supplemental figure 2). Indeed, *n7bmlΔ* strongly accumulates a Q module intermediates, similar in size to that seen in *ind1Δ*. *n7bmΔ* also accumulates Q module intermediates, as well as a higher ~500 kDa intermediate containing N, Q and membrane arm subunits, again highly reminiscent of those seen in *ind1Δ*. N7BM is a N module subunit and the N7BML human homolog NDUFAF2 has a putative role binding the N module to the Q module. The similarity of *ind1Δ* to these mutants points to a role for Ind1 in assembly of the matrix arm. Interestingly in the same paper, (Kmita et al. 2015), a role for N7BML in keeping subunits in an open state while FeS clusters were inserted was proposed. A role for the zinc binding NUMM subunit in this process was also proposed.

In humans and plant, different assembly intermediates are seen, although the overall picture is similar. Analysis of patient complex I assembly profiles (Kevelam et al. 2013) suggest there is no accumulation of the matrix arm, in contrast to what I see in *Yarrowia lipolytica*. They suggest that NUBPL play a role early in the assembly process, inserting clusters into subunits, before they are assembled into larger intermediates. Analysis in



HeLa cells, where NUBPL has been silenced, indicates an accumulation of an intermediate containing membrane subunits (Sheftel et al. 2009). Additionally, the protein levels of matrix arm subunits are decreased, whereas the membrane arm ones are relatively unaffected (Sheftel et al. 2009; Stehling et al. 2014). Similarly, in *Arabidopsis thaliana*, intermediates containing membrane subunits were found (Wydro et al. 2013). Together these data suggest a defect in the assembly of the matrix arm, where the FeS clusters are contained.

In summary, the main mutant phenotype appears to be defective formation of the matrix arm of complex I, suggesting a putative role for Ind1 as an assembly factor for the matrix arm. A role for insertion of FeS clusters cannot be shown without interaction studies.

### 6.1.3 A role for INDH in mitochondrial translation

INDH was thought to play a direct role, but a recent study in plants suggests another possibility. That study, (Wydro et al. 2013), suggests a role in mitochondrial translation. INDH mutants had lower levels of mitochondrial translation, as shown by <sup>35</sup>S-Met labelling. In addition, gametogenesis phenotypes more closely resembled the phenotypes of mutants in mitochondrial translation rather than complex I. Finally, INDH protein levels and complex I reassembly rates did not correlate, indicating a more indirect role in complex I assembly. The idea is that a defect in translation would impact the production of mitochondrially encoded complex I subunits, and so impact complex I assembly. Complex I would be worst affected of the respiratory complexes, as it has the greatest number of mitochondrially encoded subunits (9 in *Arabidopsis thaliana*). Something similar can be seen in humans, where mutations in mt-tRNAs lead to complex I deficiency, whilst the other respiratory complexes are unaffected (Da Pozzo et al. 2009; Blakely et al. 2009; Calvo et al. 2010; Swalwell et al. 2011). This shows that translation defects can affect complex I specifically, making a role for INDH in translation appear plausible. However, a mutant in the gene encoding the ribosomal protein RPS10 in *Arabidopsis thaliana* caused a decrease in mitochondrial translation, resulting in a severe decrease in complexes I, IV and V, suggesting a translation defect in plants affects all complexes (Kwasniak et al. 2013). It has been shown that if complex I is inhibited with rotenone, this can have an effect of mitochondrial translation (Allen et al. 1995). This means that complex I deficiency could be the cause, rather than the consequence, of a decrease in mitochondrial translation. Wydro et al. realised this possibility and supplemented the isolated mitochondria with ATP to overcome the presumed effect of a lack of complex I. However, complex I deficiency has other effects. First, (Allen et al. 1995) showed that a change in the redox state of the mitochondria caused by inhibition

of complex I could affect mitochondrial translation. Second, complex I significantly contributes to the PMF, which is required for import of the translation machinery. It is plausible that with a lower PMF, import is inhibited and so translation decreased. To distinguish these possibilities <sup>35</sup>S-Met labelling could be carried out on confirmed complex I mutants, to see if mitochondrial translation was affected. The 2013 study only looked at translation in a partially complemented mutant, with low, but not absent, INDH protein levels, so I carried out Western blot analysis on the full knock-out mutant, to investigate the protein levels of mitochondrially encoded subunits (Chapter 5). Contrary to expectation for a mitochondrial translation mutant, the mitochondrially encoded complex I subunit, nad9, was unaffected. In addition, the mitochondrially encoded complex IV subunit, cox2, was slightly increased. Additionally, complex IV levels were strongly increased, in contrast to that seen in (Kwasniak et al. 2013). These observations do not suggest a defect in overall mitochondria translation, but do not preclude a more specialised role for INDH in mitochondrial translation. In other species, no evidence has been found of a role in translation. In humans, there is evidence of an accumulation of membrane intermediates, but a decrease in matrix arm subunits (Sheftel et al. 2009). Given that the membrane arm contains the mitochondrially encoded subunits in humans, and the matrix arm is nuclear encoded, these data do not suggest a translation defect. Taken together these observations show that the possibility of a role for INDH in translation cannot be excluded but it is not well understood and needs more investigation. Further studies in yeast and humans will be needed to tell if a role in translation is conserved or plant specific. In summary, it is hard to say if Ind1/NUBPL/INDH is directly or indirectly involved in complex I assembly.

## 6.2 FeS provision for the respiratory complexes

### 6.2.1 Defects in the ISC pathway impact on respiratory complex assembly

Given there are still unanswered questions about the role of Ind1 in FeS cluster insertion, and the almost certainty that other factors are involved, it is a good idea to visit the question of FeS cluster delivery to respiratory complexes. Respiratory complexes, as well as all mitochondrial FeS proteins, are dependent on the ISC pathway, for example the cysteine desulphurase, Nfs1, and the scaffold protein, Isu1. Patients with mutations in the human homolog of Isu1, *ISCU1*, have defects in respiratory complexes I, II and III (Haller et al. 1991; Hall et al. 1993; Sanaker et al. 2010). Patients with mutations in *LYRM4*, the human homolog of Isd11, have defective NFS1 and so display defects in complexes I, II and III (Lim et al. 2013). The lack of respiratory complex defect seen in

the *isu1-1* mutant in *Arabidopsis* (chapter 5), is therefore surprising. However, this was only a partial mutant and so sufficient ISU1 protein may be present to support respiratory complex formation. This is lent support by the observation that a lack of ISU1 in *Arabidopsis thaliana* is fatal (Delphine Bernard and Janneke Balk, unpublished data). Additionally, the two paralogs, ISU2 and ISU3, may have redundant functions and compensate for the lack of ISU1.

The factors in the later part of the ISC pathway are less clear. Good evidence comes from patient mutations. For example, Nfu1 patients have a decrease in complex II (Navarro-Sastre et al. 2011; Cameron et al. 2011), suggesting Nfu1 is necessary for FeS cluster delivery to this respiratory complex. This defect is also seen in *Saccharomyces cerevisiae* (Schilke et al. 1999; Melber et al. 2016). A defect in complex I has also been observed in a human patient (Cameron et al. 2011). Nfu1 mutants have also been shown in yeast and human to be defective in lipoic acid cofactor containing enzymes (Navarro-Sastre et al. 2011; Cameron et al. 2011; Melber et al. 2016). In my data from *Arabidopsis thaliana*, the hemizygous *nfu4nfu5* mutant was not defective in complex I or II, but did have a mild defect in a lipoic acid containing enzyme. From these data it is unclear if Nfu1 is a specific FeS carrier for complex II and lipoic acid, or rather simply acts downstream of multiple [4Fe-4S] cluster containing proteins.

Similarly, with Grx5, there is uncertainty. In the model based on data from *Saccharomyces cerevisiae* it plays a central role, meaning a defect in Grx5 should affect the majority of FeS containing enzymes. However, in patients with mutations in the human homolog, *GLRX5*, only a subset of FeS containing enzymes are affected. Patients display anemia and a defect in lipoic acid containing enzymes (Stehling et al. 2014), but respiratory complexes were unaffected (Baker et al. 2014). In *Arabidopsis thaliana*, despite a severe growth defect there appear very few FeS defects (Stroeher et al. 2015; my data in chapter 5). The respiratory complexes appear entirely unaffected and the defects in lipoic acid containing enzymes is mild. This implies that either GRXS15 in plants plays a different role than in yeast or that it does not play as central a role in FeS cluster synthesis as is currently thought.

Studies on other ISC components can also give clues about supply of FeS clusters to the respiratory chain. Depletion of IBA57, ISCA1 or ISCA2 in HeLa cells, resulted in a decrease in complex I and complex II levels, as well as other mitochondrial [4Fe-4S] proteins (Sheftel et al. 2012). In addition, a patient with a mutation in IBA57 showed similar biochemical defects (Bolar et al. 2013). A patient with mutations in *ISCA2* displayed a defect in only complex I (Al-Hassnan et al. 2015), although the

characterisation is incomplete, and given the results in HeLa cells it is unlikely ISCA2 has a complex I specific effect. A recently reported patient with mutations in *ISCA1* (Shukla et al. 2017) similarly was reported to suffer from multiple mitochondrial dysfunction syndrome. In *Saccharomyces cerevisiae*, double mutants of *bol1/3Δ* are defective in complex II and aconitase (Melber et al. 2016; Uzarska et al. 2016). Patients with mutations in the human homolog, *BOLA3*, have defects in complex I, II, III as well as other FeS proteins (Cameron et al. 2011). Clearly, defects in ISC machinery can affect FeS provision to the respiratory complexes, although the specificity of these defects is still unclear.

### **6.2.2 LYR motif proteins involved with FeS delivery to respiratory complexes**

Recent work in mammalian FeS assembly has suggested specific factors needed for FeS insertion into the respiratory complexes. It has been proposed that a complex of HSC20-HSPA9-ISCU1 directly inserts clusters into complexes I, II and III (Maio et al. 2014; Maio et al. 2017). This was first shown for the complex II FeS subunit, SDHB, which was shown to bind the HSC20-HSPA9-ISCU1 complex (Maio et al. 2014). This interaction was dependent of a conserved LYR motif in the SDHB subunit. Notably this motif is also found in other proteins, the LYRM proteins, for example LYRM4 (human homolog of Isd11), LYRM7, needed for complex III assembly, and subunits of complex I, LYRM3 and LYRM6 (Angerer 2015). Indeed, it has been suggested the LYRM6 subunit of complex I could influence FeS incorporation (Angerer et al. 2014; Rouault & Maio 2017). Further work showed that the HSC20-HSPA9-ISCU1 complex directly binds LYRM7, which facilitates FeS transfer to the complex III FeS subunit, UQCRC1 (Maio et al. 2017). That same study showed HSC20 bound to the five FeS subunits of complex I (NDUFS1, NDUFS8, NDUFV2, NDUFV1 and NDUFS7), and suggest that the HSC20-HSPA9-ISCU1 complex directly inserts FeS clusters. Additionally, HSC20 was found to bind to numerous other FeS assembly proteins, including ISCA1, IBA57 and NFU1. The study dismisses NUBPL as a FeS insertion factor and does not mention it experimentally, however a closer look at the proteomic data shows that NUBPL bound to HSC20 in the same proportions as ISCA1 and NFU1 (2 unique peptides identified) (Maio et al. 2017), supplemental table S1). HSC20 appears to interact with all key players in FeS assembly, and care must be taken in assigning specific interactions. Notably, reciprocal co-IP was not performed to confirm the interaction between complex I subunits and HSC20. Further work is needed to confirm the interaction with the complex I subunits. It is likely that HSC20-HSPA9-ISCU1 is needed for FeS cluster incorporation into the respiratory

complexes, however further interaction studies are needed to confirm the process for complex I.

Clearly any ISC components essential for FeS generation will affect respiratory complex FeS provision. Mutants in numerous other FeS assembly proteins, for example NFU1, IBA57, ISCA1, ISCA2, NUBPL, also impact on respiratory complex levels. Recent exciting findings suggest a direct role for HSC20-HSPA9-ISCU1 in inserting FeS clusters.

### **6.3 Distinguishing between two models of FeS cluster assembly in the mitochondria**

Two models for the ISC FeS assembly pathway were presented in the introduction. One is based mainly on observations from the yeast *Saccharomyces cerevisiae*, although it does incorporate data from human models, whilst the other is based on observations from mammals, predominantly humans. The models agree on the first steps in the pathway, about FeS generation using Nfs1 and Isu1, but diverge later on the role of the FeS carrier proteins that are investigated in this thesis. Some of these differences may be due to lineage specific differences in the pathway. For example, in mammals there is evidence for a cytosolic ISC pathway, which calls into question the role of the CIA pathway (Rouault 2012).

The pathway in *Saccharomyces cerevisiae* emphasises many discrete steps, with a clearly delineated and separate “late ISC machinery” needed for [4Fe-4S] cluster delivery (Braymer & Lill 2017). The mammalian model emphasises the chaperone/co-chaperone complex around Isu1 as the major branch point. It suggests clusters are delivered from this complex, sometimes with the help of different subsets of intermediary scaffolds (Maio & Rouault 2015; Rouault & Maio 2017) (see section 6.2.2).

Looking at the data gathered in plants in chapter 5, is it possible to distinguish between these two models? The first model, from *Saccharomyces cerevisiae*, emphasises the central role of Grx5 in connecting Isu1 to the late part of the ISC assembly pathway, claiming “Grx5 and its bound FeS cluster are required for maturation of all cellular FeS proteins” (Uzarska et al. 2013). The second model de-emphasises the role of GLRX5, instead listing it as one of many FeS carriers (Maio & Rouault 2015). My data points to a very mild defect in FeS assembly in *Arabidopsis thaliana* GRXS15 mutants, in contrast to its severe growth phenotype. In plants, it is likely that GRXS15 does not play a central role in the mitochondrial ISC pathway.

Interesting lineage specific differences exist. For example, *nfu1* $\Delta$  have comparatively mild defects in *Saccharomyces cerevisiae*, but mutations in human *NFU1* can lead to a fatal disease. By studying mitochondrial NFU proteins in *Arabidopsis thaliana* a third data point can be taken, allowing us to tell which is the outlier. My data, in chapter 5, suggests that a lack of both NFU4 and NFU5 is fatal, similar to humans. This suggests that *Saccharomyces cerevisiae* is an outlier and NFU proteins are very important for FeS assembly, although their precise role remains unknown.

Constructing a general model for eukaryotic FeS assembly is hard. It needs to reconcile experimental observations from numerous study systems. Additionally, matters are complicated as lineage specific differences are present. The model needs further development based on new data.

### **6.4 *Yarrowia lipolytica* as a model for complex I assembly**

For a model organism for complex I research, *Yarrowia lipolytica* complex I biogenesis is relatively understudied. Large scale complexome sequencing to uncover the sequence of module creation, like those carried out in humans (Guerrero-Castillo et al. 2016), has not been undertaken. Instead, the process must be pieced together from observations of assembly intermediates formed in individual mutants, for example in NUMM, N7BM, N7BML and NUYM (Kmita et al. 2015; Kahlhöfer et al. 2017) and my own data, chapter 3. The process of *Yarrowia lipolytica* complex I assembly can be recreated as follows: The N and the Q module are formed separately and then are joined to form the matrix arm. These then either attach with the full membrane arm, to form mature complex I, or associate with part of the membrane arm before, the remaining membrane arm subunits are added, forming mature complex I.

It has been presumed that *Yarrowia lipolytica* complex I assembly would be similar to that seen in another model yeast, *Neurospora crassa*, however, there has recently been uncertainty about this process. In this case, the process in *Yarrowia lipolytica* may prove instructive for *Neurospora crassa* complex I assembly, rather than *vice versa*. It has long been thought that in the matrix and membrane arms form independently, and then join together to form mature complex I, in contrast to mammals, where the N module is added last (Videira & Duarte 2002; Vogel et al. 2007). However, recently it has been proposed that the process in *Neurospora crassa* is similar to mammals and the N module is added last (Pereira et al. 2013). My data suggest that the N and Q module form the matrix arm first, before associating with the membrane arm, lending support to the first model.

Only two assembly factors have been studied in *Yarrowia lipolytica*: Ind1 (human NUBPL) (Bych et al. 2008) and N7BML (human NDUFAF2) (Kmita et al. 2015). Homologs exist in *Yarrowia lipolytica* for 8 out of 16 assembly factors found in humans. Many of these will presumably play a similar role to the human version. These homologs have yet to be functionally characterised. It is likely there are many, yeast specific, assembly factors in *Yarrowia lipolytica* waiting to be found, to fulfil the functions of the assembly factors for which there are no homologs.

*Yarrowia lipolytica* has been used numerous times to study human complex I deficiency, usually looking at structural subunits (Ahlers et al. 2000; Kerscher et al. 2004; Varghese et al. 2015; Gerber et al. 2016). However, if it is to also be used to study complex I assembly deficiency, as it has in this thesis, the process needs to be understood better. However, given the differences in assembly between the two groups, *Yarrowia lipolytica* may be of limited use in studying human defect in complex I assembly in the future.

## 7 Bibliography

- Abdrakhmanova, A. et al., 2005. Functional sulfurtransferase is associated with mitochondrial complex I from *Yarrowia lipolytica*, but is not required for assembly of its iron-sulfur clusters. *FEBS Letters*, 579(30), pp.6781–6785.
- Adams, K.L. et al., 2002. Punctuated evolution of mitochondrial gene content: high and variable rates of mitochondrial gene loss and transfer to the nucleus during angiosperm evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 99(15), pp.9905–9912.
- Adzhubei, I.A. et al., 2010. A method and server for predicting damaging missense mutations. *Nature Methods*, 7(4), pp.248–249.
- Agar, J.N. et al., 2000. Modular organization and identification of a mononuclear iron-binding site within the NifU protein. *Journal of Inorganic Biochemistry*, 5, pp.167–177.
- Ahlers, P.M. et al., 2000. Application of the obligate aerobic yeast *Yarrowia lipolytica* as a eucaryotic model to analyse Leigh syndrome mutations in the complex I core subunits PSST and TYKY. *Biochimica et Biophysica Acta - Bioenergetics*, 1459(2–3), pp.258–265.
- Ahting, U. et al., 2015. Clinical, biochemical, and genetic spectrum of seven patients with NFU1 deficiency. *Frontiers in Genetics*, 6(April), pp.1-13.
- Al-Hassnan, Z.N. et al., 2015. ISCA2 mutation causes infantile neurodegenerative mitochondrial disorder. *Journal of Medical Genetics*, 52, pp.186–194.
- Allen, C., Hakansson, G. & Allen, J., 1995. Redox conditions specify the proteins synthesised by isolated chloroplasts and mitochondria. *Redox Report*, 1, pp.119–123.
- Alston, C.L. et al., 2017. The genetics and pathology of mitochondrial disease. *Journal of Pathology*, 241, pp.236–250.
- Anderson, S. et al., 1981. Sequence and organization of the human mitochondrial genome. *Nature*, 290(5806), pp.457–465.
- Andrews, B. et al., 2013. Assembly factors for the membrane arm of human complex I. *Proceedings of the National Academy of Sciences of the United States of America*, 110(47), pp.18934–18939.



Andriotis, V.M. et al., 2010. The plastidial glucose-6-phosphate/phosphate antiporter GPT1 is essential for morphogenesis in Arabidopsis embryos. *Plant Journal*, 64(1), pp.128–139.

Angerer, H. et al., 2011. A scaffold of accessory subunits links the peripheral arm and the distal proton-pumping module of mitochondrial complex I. *Biochemical Journal*, 437(2), pp.279–288.

Angerer, H., 2015. Eukaryotic LYR Proteins Interact with Mitochondrial Protein Complexes. *Biology*, 4(1), pp.133–150.

Angerer, H. et al., 2014. The LYR protein subunit NB4M/NDUFA6 of mitochondrial complex I anchors an acyl carrier protein and is essential for catalytic activity. *Proceedings of the National Academy of Sciences of the United States of America*, 111(14), pp.5207–5212.

Antonicka, H. et al., 2003. Identification and Characterization of a Common Set of Complex I Assembly Intermediates in Mitochondria from Patients with Complex I Deficiency. *Journal of Biological Chemistry*, 278(44), pp.43081–43088.

Arroyo, J.D. et al., 2016. A Genome-wide CRISPR Death Screen Identifies Genes Essential for Oxidative Phosphorylation. *Cell Metabolism*, 24(6), pp.875–885.

Baker, P.R. et al., 2014. Variant non ketotic hyperglycinemia is caused by mutations in LIAS, BOLA3 and the novel gene GLRX5. *Brain*, 137(2), pp.366–379.

Balk, J. & Schaedler, T. A, 2014. Iron cofactor assembly in plants. *Annual Review of Plant Biology*, 65, pp.125–153.

Balsa, E. et al., 2012. NDUFA4 is a subunit of complex IV of the mammalian electron transport chain. *Cell Metabolism*, 16(3), pp.378–386.

Bandyopadhyay, S. et al., 2008. Chloroplast monothiol glutaredoxins as scaffold proteins for the assembly and delivery of [2Fe-2S] clusters. *The EMBO journal*, 27(7), pp.1122–1133.

Baradaran, R. et al., 2013. Crystal structure of the entire respiratory complex I. *Nature*, 494(7438), pp.443–8.

Bastow, E.L. et al., 2017. NBP35 interacts with DRE2 in the maturation of cytosolic iron-sulfur proteins in Arabidopsis thaliana. *The Plant Journal*, pp.590–600.

## Chapter 7 Bibliography

- Beinert, H., 2000. Iron-sulfur proteins: ancient structures, still full of surprises. *Journal of Biological Inorganic Chemistry*, 5(1), pp.2–15.
- Beinert, H., Holm, R.H. & Münck, M., 1997. Iron-Sulfur Clusters: Nature's Modular, Multipurpose Structures. *Science*, 277(August), p.653-659.
- Bell, C.J. et al., 2011. Carrier Testing for Severe Childhood Recessive Diseases by Next-Generation Sequencing. *Science Translational Medicine*, 3(65), pp.1–26.
- Benz, C. et al., 2016. Roles of the Nfu Fe-S targeting factors in the trypanosome mitochondrion. *International Journal for Parasitology*, 46, pp.641–651.
- Bernard, D.G. et al., 2009. An allelic mutant series of ATM3 reveals its key role in the biogenesis of cytosolic iron-sulfur proteins in Arabidopsis. *Plant Physiology*, 151(2), pp.590–602.
- Bernard, D.G. et al., 2013. Requirements of the cytosolic iron-sulfur cluster assembly pathway in Arabidopsis. *Philosophical Transactions Royal Society of London. Series B, Biological Sciences*, 368(1622).
- Berrisford, J.M., Baradaran, R. & Sazanov, L.A., 2016. Structure of bacterial respiratory complex I. *Biochimica et Biophysica Acta - Bioenergetics*, 1857(7), pp.4–13.
- Björkholm, P. et al., 2015. Mitochondrial genomes are retained by selective constraints on protein targeting. *Proceedings of the National Academy of Sciences of the United States of America*, 112(33), pp. 10154-10161
- Blakely, E.L. et al., 2009. A New Mitochondrial Transfer RNA. *Archives of Neurology*, 66(3), pp.399–402.
- Bolar, N.A. et al., 2013. Mutation of the iron-sulfur cluster assembly gene IBA57 causes severe myopathy and encephalopathy. *Human Molecular Genetics*, 22(13), pp.2590–2602.
- Bourgeron, T. et al., 1995. Mutation of a nuclear succinate dehydrogenase gene results in mitochondrial respiratory chain deficiency. *Nature Genetics*, 11(2), pp.144–149.
- Brandt, U., 2006. Energy converting NADH:quinone oxidoreductase (complex I). *Annual Review of Biochemistry*, 75(Complex I), pp.69–92.
- Braymer, J.J. & Lill, R., 2017. Iron-Sulfur Cluster Biogenesis and Trafficking in Mitochondria. *Journal of Biological Chemistry*, 292, pp.12754-12763

- Bugiani, M. et al., 2004. Clinical and molecular findings in children with complex I deficiency. *Biochimica et Biophysica Acta*, 1659(2–3), pp.136–147.
- Burki, F., 2016. Mitochondrial Evolution: Going, Going, Gone. *Current Biology*, 26(10), pp.R410–R412.
- Busi, M. V. et al., 2006. Deficiency of *Arabidopsis thaliana* frataxin alters activity of mitochondrial Fe-S proteins and induces oxidative stress. *Plant Journal*, 48(6), pp.873–882.
- Bych, K. et al., 2008. The iron-sulphur protein Ind1 is required for effective complex I assembly. *EMBO Journal*, 27(12), pp.1736–1746.
- Bych, K. et al., 2008b. The essential cytosolic iron-sulfur protein Nbp35 acts without Cfd1 partner in the green lineage. *The Journal of Biological Chemistry*, 283(51), pp.35797–35804.
- Calvo, S.E. et al., 2010. High-throughput, pooled sequencing identifies mutations in NUBPL and FOXRED1 in human complex I deficiency. *Nature Genetics*, 42(10), pp.851–858.
- Calvo, S.E., Clauser, K.R. & Mootha, V.K., 2016. MitoCarta2.0: An updated inventory of mammalian mitochondrial proteins. *Nucleic Acids Research*, 44(D1), pp.D1251–D1257.
- Calvo, S.E. & Mootha, V.K., 2010. The mitochondrial proteome and human disease. *Annual Review of Genomics and Human Genetics*, 11, pp.25–44.
- Cameron, J.M. et al., 2011. Mutations in iron-sulfur cluster scaffold genes NFU1 and BOLA3 cause a fatal deficiency of multiple respiratory chain and 2-oxoacid dehydrogenase enzymes. *American Journal of Human Genetics*, 89(4), pp.486–495.
- Carilla-Latorre, S. et al., 2010. MidA is a putative methyltransferase that is required for mitochondrial complex I function. *Journal of Cell Science*, 123, pp.1674–1683.
- Carroll, J. et al., 2006. Bovine complex I is a complex of 45 different subunits. *Journal of Biological Chemistry*, 281(43), pp.32724–32727.
- Cavazza, C. et al., 2008. The possible role of an [FeFe]-hydrogenase-like protein in the plant responses to changing atmospheric oxygen levels. *Journal of Inorganic Biochemistry*, 102(5–6), pp.1359–1365.

Chen, D.C., Beckerich, J.M. & Gaillardin, C., 1997. One-step transformation of the dimorphic yeast *Yarrowia lipolytica*. *Applied Microbiology and Biotechnology*, 48(2), pp.232–235.

Cheng, N.H., 2008. AtGRX4, an Arabidopsis chloroplastic monothiol glutaredoxin, is able to suppress yeast grx5 mutant phenotypes and respond to oxidative stress. *FEBS Letters*, 582(6), pp.848–854.

Clark, C.G. & Rogert, A.J., 1995. Direct evidence for secondary loss of mitochondria in *Entamoeba histolytica*. *Proceedings of the National Academy of Sciences of the United States of America*, 92(14), pp.6518–6521.

Couturier, J. et al., 2014. Monothiol glutaredoxin-BolA interactions: Redox control of arabidopsis thaliana BolA2 and SufE1. *Molecular Plant*, 7(1), pp.187–205.

Couturier, J. et al., 2013. The iron-sulfur cluster assembly machineries in plants: current knowledge and open questions. *Frontiers in Plant Science*, 4(July), pp.1-22.

Craven, L. et al., 2017. Recent Advances in Mitochondrial Disease. *Annual Review of Genomics and Human Genetics*, 18(1), pp.257-275

Dörner, K. et al., 2017. Significance of [2Fe-2S] Cluster N1a for Electron Transfer and Assembly of Escherichia coli Respiratory Complex I. *Biochemistry*, 56(22), pp.2770–2778.

Duarte, M. et al., 2005. Neurospora Strains Harboring Mitochondrial Disease-Associated Mutations in Iron-Sulfur Subunits of Complex I. *Genetics*, 171(1), pp.91–99.

Dudkina, N. V et al., 2005. Structure of a mitochondrial supercomplex formed by respiratory-chain complexes I and III. *Proceedings of the National Academy of Sciences of the United States of America*, 102(9), pp.3225–3229.

Edwards, K., Johnstone, C. & Thompson, C., 1991. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Research*, 19(6), p.1349.

Embley, T.M. & Martin, W., 2006. Eukaryotic evolution, changes and challenges. *Nature*, 440(7084), pp.623–630.

Eubel, H., Jansch, L., & Braun, H., 2003. New Insights into the Respiratory Chain of Plant Mitochondria. Supercomplexes and a Unique Composition of Complex II. *Plant Physiology*, 133(September), pp.274–286.

- Fassone, E. et al., 2010. FOXRED1, encoding an FAD-dependent oxidoreductase complex-I-specific molecular chaperone, is mutated in infantile-onset mitochondrial encephalopathy. *Human Molecular Genetics*, 19(24), pp.4837–4847.
- Fassone, E. & Rahman, S., 2012. Complex I deficiency: clinical features, biochemistry and molecular genetics. *Journal of Medical Genetics*, 49(9), pp.578–590.
- Feissner, R., Xiang, Y. & Kranz, R.G., 2003. Chemiluminescent-based methods to detect subpicomole levels of c-type cytochromes. *Analytical Biochemistry*, 315(1), pp.90–94.
- Ferreira, M. et al., 2011. Progressive cavitating leukoencephalopathy associated with respiratory chain complex I deficiency and a novel mutation in NDUF51. *Neurogenetics*, 12(1), pp.9–17.
- Ferrer-Cortès, X. et al., 2013. Protein expression profiles in patients carrying NFU1 mutations. Contribution to the pathophysiology of the disease. *Journal of Inherited Metabolic Disease*, 36, pp.841–847.
- Fiedorczuk, K. et al., 2016. Atomic structure of the entire mammalian mitochondrial complex I. *Nature*, 538(7625), pp.1–21.
- Flegonotov, P. et al., 2015. Divergent Mitochondrial Respiratory Chains in Phototrophic Relatives of Apicomplexan Parasites. *Molecular Biology and Evolution*, 32(5), pp.1115–1131.
- Floyd, B.J. et al., 2016. Mitochondrial Protein Interaction Mapping Identifies Regulators of Respiratory Chain Function. *Molecular Cell*, 63(4), pp.621–632.
- Formosa, L.E. et al., 2017. Building a complex complex: Assembly of mitochondrial respiratory chain complex I. *Seminars in Cell & Developmental Biology*, in press.
- Formosa, L.E. et al., 2015. Characterization of mitochondrial FOXRED1 in the assembly of respiratory chain complex I. *Human Molecular Genetics*, 24(10), pp.2952–2965.
- Frazzon, A.P. et al., 2007. Functional analysis of Arabidopsis genes involved in mitochondrial iron-sulfur cluster assembly. *Plant Molecular Biology*, 64(3), pp.225–240.
- Fromm, S. et al., 2016. Life without complex I: proteome analyses of an Arabidopsis mutant lacking the mitochondrial NADH dehydrogenase complex. *Journal of Experimental Botany*, 67(10), pp.3079–3093.

Fromm, S. et al., 2016b. The carbonic anhydrase domain of plant mitochondrial complex I. *Physiologia Plantarum*, 157, pp.289–296.

Gabaldón, T., Rainey, D. & Huynen, M.A., 2005. Tracing the Evolution of a Large Protein Complex in the Eukaryotes, NADH: Ubiquinone Oxidoreductase (Complex I). *Journal of Molecular Biology*, 348, pp.857–870.

Gerber, S. et al., 2016. Compound heterozygosity for severe and hypomorphic NDUF52 mutations cause non-syndromic LHON-like optic neuropathy. *Journal of Medical Genetics*, pp.1–11.

Gorman, G.S. et al., 2015. Prevalence of nuclear and mitochondrial DNA mutations related to adult mitochondrial disease. *Annals of Neurology*, 77(5), pp.753–759.

Gray, M.W., Burger, G. & Lang, B.F., 1999. Mitochondrial Evolution. *Science*, 283(5407), pp.1–16.

Guerrero-Castillo, S. et al., 2016. The Assembly Pathway of Mitochondrial Respiratory Chain Complex I. *Cell Metabolism*, 25(1), pp.128-139

Haack, T.B. et al., 2010. Exome sequencing identifies ACAD9 mutations as a cause of complex I deficiency. *Nature Genetics*, 42(12), pp.1131–1134.

Hall, R.E. et al., 1993. Mitochondrial myopathy with succinate dehydrogenase and aconitase deficiency abnormalities of several iron-sulfur proteins. *Journal of Clinical Investigation*, 92(6), pp.2660–2666.

Haller, R.G. et al., 1991. Deficiency of skeletal muscle succinate dehydrogenase and aconitase. Pathophysiology of exercise in a novel human muscle oxidative defect. *Journal of Clinical Investigation*, 88(4), pp.1197–1206.

Heide, H. et al., 2012. Complexome profiling identifies TMEM126B as a component of the mitochondrial complex I assembly complex. *Cell Metabolism*, 16(4), pp.538–549.

Hirst, J., 2013. Mitochondrial complex I. *Annual Review of Biochemistry*, 82, pp.551–575.

Hjort, K. et al., 2010. Diversity and reductive evolution of mitochondria among microbial eukaryotes. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 365(1541), pp.713–727.

- Hjorth, E. et al., 2005. Unique genetic compartmentalization of the SUF system in cryptophytes and characterization of a SufD mutant in *Arabidopsis thaliana*. *FEBS Letters*, 579(5), pp.1129–1135.
- Holt, I.J., Harding, A.E. & Morgan-Hughes, J.A., 1988. Deletions of muscle mitochondrial DNA in patients with mitochondrial myopathies. *Nature*, 331, pp.717–719.
- Hsu, Y.W. et al., 2014. *Arabidopsis* mTERF15 is required for mitochondrial nad2 intron 3 splicing and functional complex I activity. *PLoS ONE*, 9(11), pp.e112360
- Huang, M. et al., 2013. Construction of plastid reference proteomes for maize and *Arabidopsis* and evaluation of their orthologous relationships; The concept of orthoproteomics. *Journal of Proteome Research*, 12(1), pp.491–504.
- Hunte, C., Zickermann, V. & Brandt, U., 2010. Functional modules and structural basis of conformational coupling in mitochondrial complex I. *Science*, 329(5990), pp.448–451.
- Invernizzi, F. et al., 2014. Cavitating leukoencephalopathy with multiple mitochondrial dysfunction syndrome and NFU1 mutations. *Frontiers in Genetics*, 5(November), p.412.
- James, T.Y. et al., 2013. Shared signatures of parasitism and phylogenomics unite cryptomycota and microsporidia. *Current Biology*, 23(16), pp.1548–1553.
- Janssen, R. et al., 2002. CIA30 complex I assembly factor: A candidate for human complex I deficiency? *Human Genetics*, 110(3), pp.264–270.
- Jones, A.J.Y. et al., 2017. Respiratory complex I in *bos Taurus* and *Paracoccus denitrificans* pumps four protons across the membrane for every NADH oxidized. *Journal of Biological Chemistry*, 292(12), pp.4987–4995.
- Kahlhöfer, F. et al., 2017. Accessory subunit NUYM (NDUFS4) is required for stability of the electron input module and activity of mitochondrial complex I. *Biochimica et Biophysica Acta - Bioenergetics*, 1858(2), pp.175–181.
- Karnkowska, A. et al., 2016. A Eukaryote without a Mitochondrial Organelle. *Current Biology*, pp.1274–1284.
- Kasajima, I. et al., 2004. A protocol for rapid DNA extraction from *Arabidopsis thaliana* for PCR analysis. *Plant Molecular Biology Reporter*, 22(1), pp.49–52.

Kerscher, S. et al., 2004. Application of the yeast *Yarrowia lipolytica* as a model to analyse human pathogenic mutations in mitochondrial complex I (NADH:ubiquinone oxidoreductase). *Biochimica et Biophysica Acta*, 1659(2–3), pp.197–205.

Kerscher, S. et al., 2002. *Yarrowia lipolytica*, a yeast genetic system to study mitochondrial complex I. *Biochimica et Biophysica Acta*, 1555(1–3), pp.83–91.

Kerscher, S.J. et al., 2001. External alternative NADH:ubiquinone oxidoreductase redirected to the internal face of the mitochondrial inner membrane rescues complex I deficiency in *Yarrowia lipolytica*. *Journal of Cell Science*, 114, pp.3915–3921.

Kerscher, S.J., Okun, J.G. & Brandt, U., 1999. A single external enzyme confers alternative NADH:ubiquinone oxidoreductase activity in *Yarrowia lipolytica*. *Journal of Cell Science*, 112, pp.2347–2354.

Kevelam, S.H., Rodenburg, R.J., et al., 2013. NUBPL mutations in patients with complex I deficiency and a distinct MRI pattern. *Neurology*, 80(17), pp.1577–1583.

Klodmann, J. et al., 2011. Defining the Protein Complex Proteome of Plant Mitochondria. *Plant Physiology*, 157(October), pp.587–598.

Kmita, K. et al., 2015. Accessory NUMM (NDUFS6) subunit harbors a Zn-binding site and is essential for biogenesis of mitochondrial complex I. *Proceedings of the National Academy of Sciences of the United States of America*, 112(18), pp.5685–5690

Kmita, K. & Zickermann, V., 2013. Accessory subunits of mitochondrial complex I. *Biochemical Society Transactions*, 41, pp.1272–1279.

Koene, S. et al., 2012. Natural disease course and genotype-phenotype correlations in Complex I deficiency caused by nuclear gene defects: What we learned from 130 cases. *Journal of Inherited Metabolic Disease*, 35(5), pp.737–747.

Kohbushi, H. et al., 2009. Arabidopsis cytosolic Nbp35 homodimer can assemble both [2Fe-2S] and [4Fe-4S] clusters in two distinct domains. *Biochemical and Biophysical Research Communications*, 378(4), pp.810–815.

Küffner, R. et al., 1998. Involvement of two novel chaperones in the assembly of mitochondrial NADH:Ubiquinone oxidoreductase (complex I). *Journal of Molecular Biology*, 283(2), pp.409–417.



Kühn, K. et al., 2011. The RCC1 family protein RUG3 is required for splicing of nad2 and complex I biogenesis in mitochondria of *Arabidopsis thaliana*. *The Plant Journal*, 67(6), pp.1067–1080.

Kwasniak, M. et al., 2013. Silencing of the nuclear RPS10 gene encoding mitochondrial ribosomal protein alters translation in *Arabidopsis* mitochondria. *The Plant Cell*, 25(5), pp.1855–1867.

Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227, pp.680–685.

Lane, N., 2005. Power, sex, suicide: Mitochondria and the meaning of life, Oxford: Oxford University Press.

Langston, J.W. et al., 1983. Chronic Parkinsonism in humans due to a product of meperidine-analog synthesis. *Science*, 219(4587), pp.979–980.

Lebigot, E. et al., 2017. Impact of mutations within the [Fe-S] cluster or the lipoic acid biosynthesis pathways on mitochondrial protein expression profiles in fibroblasts from patients. *Molecular Genetics and Metabolism*, (June).

León, G., Holuigue, L. & Jordana, X., 2007. Mitochondrial complex II is essential for gametophyte development in *Arabidopsis*. *Plant Physiology*, 143(4), pp.1534–1546.

Léon, S. et al., 2003. Iron-sulphur cluster assembly in plants: distinct NFU proteins in mitochondria and plastids from *Arabidopsis thaliana*. *The Biochemical Journal*, 371, pp.823–830.

Léon, S. et al., 2005. Mitochondrial localization of *Arabidopsis thaliana* Isu Fe-S scaffold proteins. *FEBS Letters*, 579(9), pp.1930–1934.

Letts, J.A. & Sazanov, L.A., 2015. Gaining mass: The structure of respiratory complex I from bacterial towards mitochondrial versions. *Current Opinion in Structural Biology*, 33, pp.135–145.

Lill, R., 2009. Function and biogenesis of iron-sulphur proteins. *Nature*, 460(7257), pp.831–838.

Lill, R. & Muhlenhoff, U., 2008. Maturation of iron-sulfur proteins in eukaryotes: mechanisms, connected processes, and diseases. *Annual Review of Biochemistry*, 77, pp.669–700.

- Lillig, C.H. et al., 2005. Characterization of human glutaredoxin 2 as iron-sulfur protein: a possible role as redox sensor. *Proceedings of the National Academy of Sciences of the United States of America*, 102(23), pp.8168–8173.
- Lim, S.C. et al., 2013. Mutations in LYRM4, encoding iron-sulfur cluster biogenesis factor ISD11, cause deficiency of multiple respiratory chain complexes. *Human Molecular Genetics*, 22(22), pp.4460–4473.
- Liu, C.M. & Meinke, D.W., 1998. The titan mutants of Arabidopsis are disrupted in mitosis and cell cycle control during seed development. *The Plant Journal*, 16(1), pp.21–31.
- Loeffen, J. et al., 1998. The first nuclear-encoded complex I mutation in a patient with Leigh syndrome. *American Journal of Human Genetics*, 63, pp.1598–1608.
- Loeffen, J.L. et al., 2000. Isolated complex I deficiency in children: clinical, biochemical and genetic aspects. *Human Mutation*, 15(2), pp.123–134.
- Luo, D. et al., 2012. The DUF59 family gene AE7 acts in the cytosolic iron-sulfur cluster assembly pathway to maintain nuclear genome integrity in Arabidopsis. *Plant Cell*, 24(10), pp.4135–4148.
- Maio, N. et al., 2017. A Single Adaptable Cochaperone-Scaffold Complex Delivers Nascent Iron-Sulfur Clusters to Mammalian Respiratory Chain Complexes I–III. *Cell Metabolism*, 25(4), p.945–953.
- Maio, N. et al., 2014. Cochaperone binding to LYR motifs confers specificity of iron sulfur cluster delivery. *Cell Metabolism*, 19(3), pp.445–457.
- Maio, N. & Rouault, T.A., 2015. Iron –sulfur cluster biogenesis in mammalian cells: New insights into the molecular mechanisms of cluster delivery. *Biochimica et Biophysica Acta - Molecular Cell Research*, 1853(6), pp.1493–1512.
- Margulis, L., 1970. *Origin of Eukaryotic Cells*, New Haven: Yale University Press.
- Margulis, L., 1981. *Symbiosis in Cell Evolution*, San Francisco: Freeman.
- Mathiasen, R.L. et al., 2008. Daniel L. Nickrent David M. Watson What Is a Mistletoe? *Plant Disease*, 92(7), pp.988-1006.
- Mathiesen, C. & Hägerhäll, C., 2002. Transmembrane topology of the NuoL, M and N subunits of NADH:quinone oxidoreductase and their homologues among membrane-

bound hydrogenases and bona fide antiporters. *Biochimica et Biophysica Acta - Bioenergetics*, 1556(2–3), pp.121–132.

McGuffin, L.J. et al., 2015. IntFOLD: An integrated server for modelling protein structures and functions from amino acid sequences. *Nucleic Acids Research*, 43(W1), pp.W169–W173.

McNicholas, S. et al., 2011. Presenting your structures: The CCP4mg molecular-graphics software. *Acta Crystallographica Section D: Biological Crystallography*, 67(4), pp.386–394.

Melber, A. et al., 2016. Role of Nfu1 and Bol3 in iron-sulfur cluster transfer to mitochondrial clients. *eLife*, 5, pp.e15991.

Meyer, E.H. et al., 2009. Remodeled respiration in *ndufs4* with low phosphorylation efficiency suppresses *Arabidopsis* germination and growth and alters control of metabolism at night. *Plant Physiology*, 151(2), pp.603–619.

Mimaki, M. et al., 2012. Understanding mitochondrial complex I assembly in health and disease. *Biochimica et Biophysica Acta - Bioenergetics*, 1817(6), pp.851–862.

Molina, J. et al., 2014. Possible loss of the chloroplast genome in the parasitic flowering plant *Rafflesia lagascae* (Rafflesiaceae). *Molecular Biology and Evolution*, 31(4), pp.793–803.

Moparthi, V.K. et al., 2011. Homologous protein subunits from *Escherichia coli* NADH:quinone oxidoreductase can functionally replace MrpA and MrpD in *Bacillus subtilis*. *Biochimica et Biophysica Acta - Bioenergetics*, 1807(4), pp.427–436.

Moseler, A. et al., 2015. The mitochondrial monothiol glutaredoxin S15 is essential for iron-sulfur protein maturation in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America*, 112(44), pp.13735–13740

Mower, J.P., Sloan, D.B. & Alverson, A.J., 2012. Plant Mitochondrial Genome Diversity: The Genomics Revolution, Vienna: Springer-Verlag.

Narayana Murthy, U.M. et al., 2007. Characterization of *Arabidopsis thaliana* SufE2 and SufE3: Functions in chloroplast iron-sulfur cluster assembly and NAD synthesis. *Journal of Biological Chemistry*, 282(25), pp.18254–18264.

- Navarro-Sastre, A. et al., 2011. A fatal mitochondrial disease is associated with defective NFU1 function in the maturation of a subset of mitochondrial Fe-S proteins. *American Journal of Human Genetics*, 89(5), pp.656–667.
- Nizon, M. et al., 2014. Leukoencephalopathy with cysts and hyperglycinemia may result from NFU1 deficiency. *Mitochondrion*, 15(1), pp.59–64.
- Nouws, J. et al., 2010. Acyl-CoA dehydrogenase 9 is required for the biogenesis of oxidative phosphorylation complex I. *Cell Metabolism*, 12(3), pp.283–294.
- Nouws, J. et al., 2012. Assembly factors as a new class of disease genes for mitochondrial complex I deficiency: Cause, pathology and treatment options. *Brain*, 135(1), pp.12–22.
- Nugent, J.M. & Palmer, J.D., 1991. RNA-Mediated Transfer of the Gene *coxII* from the Mitochondrion to the Nucleus during Flowering Plant Evolution. *Cell*, 66, pp.473–481.
- Ogilvie, I., Kennaway, N.G. & Shoubridge, E.A., 2005. A molecular chaperone for mitochondrial complex I assembly is mutated in a progressive encephalopathy. *Journal of Clinical Investigation*, 115(10), pp.2784–2792.
- Ohnishi, T., 1998. Iron-sulfur clusters/semiquinones in Complex I. *Biochimica et Biophysica Acta - Bioenergetics*, 1364(2), pp.186–206.
- Pagliarini, D.J. et al., 2008. A mitochondrial protein compendium elucidates complex I disease biology. *Cell*, 134(1), pp.112–123.
- Pagliarini, D.J. & Rutter, J., 2013. Hallmarks of a new era in mitochondrial biochemistry. *Genes and Development*, 27(24), pp.2615–2627.
- Pereira, B., Videira, A. & Duarte, M., 2013. Novel insights into the role of *Neurospora crassa* NDUFAF2, an evolutionarily conserved mitochondrial complex I assembly factor. *Molecular and Cellular Biology*, 33(13), pp.2623–2634.
- Petersen, G. et al., 2015. Massive gene loss in mistletoe (*Viscum*, Viscaceae) mitochondria. *Scientific Reports*, 5, pp.e17588.
- Petersen, G., Cuenca, A. & Seberg, O., 2015b. Plastome evolution in hemiparasitic mistletoes. *Genome Biology and Evolution*, 7(9), pp.2520–2532.
- Pettersen, E.F. et al., 2004. UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25(13), pp.1605–1612.

- Pitceathly, R.D.S. et al., 2013. NDUFA4 Mutations Underlie Dysfunction of a Cytochrome c Oxidase Subunit Linked to Human Neurological Disease. *Cell Reports*, 3(6), pp.1795–1805.
- Pittis, A.A. & Gabaldón, T., 2016. Late acquisition of mitochondria by a host with chimaeric prokaryotic ancestry. *Nature*, 531(7592), pp.101–104.
- Poole, A.M. et al., 2014. Eukaryotic Origins: How and When Was the Mitochondrion Acquired? *Cold Spring Harbor Perspectives in Biology*, 6(12), pp.a015990
- Poulin, R. & Morand, S., 2000. The diversity of parasites. *Quarterly Review of Biology*, 75, pp.277–293.
- Da Pozzo, P. et al., 2009. A novel mutation in the mitochondrial tRNA(Pro) gene associated with late-onset ataxia, retinitis pigmentosa, deafness, leukoencephalopathy and complex I deficiency. *European Journal of Human Genetics: EJHG*, 17(8), pp.1092–1096.
- Rasmusson, A.G., Geisler, D.A. & Møller, I.M., 2008. The multiplicity of dehydrogenases in the electron transport chain of plant mitochondria. *Mitochondrion*, 8(1), pp.47–60.
- Rhein, V.F. et al., 2016. NDUFAF5 hydroxylates NDUF57 at an early stage in the assembly of human complex I. *Journal of Biological Chemistry*, 291(28), pp.14851–14860.
- Rodenburg, R.J., 2016. Mitochondrial complex I-linked disease. *Biochimica et Biophysica Acta - Bioenergetics*, 1857(7), pp.938–945.
- Rodríguez-Manzanique, M.T. et al., 2002. Grx5 Is a Mitochondrial Glutaredoxin Required for the Activity of Iron/Sulfur Enzymes. *Molecular Biology of the Cell*, 13, pp.1109–1121.
- Rouault, T.A., 2012. Biogenesis of iron-sulfur clusters in mammalian cells: new insights and relevance to human disease. *Disease Models & Mechanisms*, 5, pp.155–164.
- Rouault, T.A., 2015. Iron-sulfur proteins hiding in plain sight. *Nature Chemical Biology*, 11(7), pp.442–445.
- Rouault, T.A. & Maio, N., 2017. Biogenesis and functions of mammalian iron-sulfur proteins in the regulation of iron homeostasis and pivotal metabolic pathways. *Journal of Biological Chemistry*, 292(31), pp.12744–12753.

Russell, M.J. & Martin, W., 2004. The rocky roots of the acetyl-CoA pathway. *Trends in Biochemical Sciences*, 29(7), pp.358–363.

Saada, A. et al., 2008. C6ORF66 Is an Assembly Factor of Mitochondrial Complex I. *American Journal of Human Genetics*, 82(1), pp.32–38.

Saada, A. et al., 2009. Mutations in NDUFAF3 (C3ORF60), Encoding an NDUFAF4 (C6ORF66)-Interacting Complex I Assembly Protein, Cause Fatal Neonatal Mitochondrial Disease. *American Journal of Human Genetics*, 84(6), pp.718–727.

Sabar, M., Balk, J. & Leaver, C.J., 2005. Histochemical staining and quantification of plant mitochondrial respiratory chain complexes using blue-native polyacrylamide gel electrophoresis. *Plant Journal*, 44(5), pp.893–901.

Salvato, F. et al., 2014. The potato tuber mitochondrial proteome. *Plant Physiology*, 164(2), pp.637–53.

Sanaker, P.S. et al., 2010. Differences in RNA processing underlie the tissue specific phenotype of ISCU myopathy. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, 1802(6), pp.539–544.

Sánchez-Caballero, L., Guerrero-Castillo, S. & Nijtmans, L., 2016. Unraveling the complexity of mitochondrial complex I assembly; a dynamic process. *Biochimica et Biophysica Acta - Bioenergetics*, 1857(7), pp.980–990.

Schaedler, T.A. et al., 2014. A conserved mitochondrial ATP-binding cassette transporter exports glutathione polysulfide for cytosolic metal cofactor assembly. *Journal of Biological Chemistry*, 289(34), pp.23264–23274.

Schilke, B. et al., 1999. Evidence for a conserved system for iron metabolism in the mitochondria of *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America*, 96(18), pp.10206–10211.

Senkler, J. et al., 2016. The mitochondrial complexome of *Arabidopsis thaliana*. *The Plant Journal*, pp.1079–1092.

Sheftel, A.D. et al., 2009. Human ind1, an iron-sulfur cluster assembly factor for respiratory complex I. *Molecular and Cellular Biology*, 29(22), pp.6059–6073.

Sheftel, A.D. et al., 2012. The human mitochondrial ISCA1, ISCA2, and IBA57 proteins are required for [4Fe-4S] protein maturation. *Molecular Biology of the Cell*, 23(7), pp.1157–1166.

Shukla, A. et al., 2017. Homozygous p.(Glu87Lys) variant in ISCA1 is associated with a multiple mitochondrial dysfunctions syndrome. *Journal of Human Genetics*, 62, pp.723–727.

Sievers, F. et al., 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7(1), pp.1-6

Skippington, E. et al., 2017. Comparative mitogenomics indicates respiratory competence in parasitic *Viscum* despite loss of complex I and extreme sequence divergence, and reveals horizontal gene transfer and remarkable variation in genome size. *BMC Plant Biology*, 17(1), pp.1-12.

Skippington, E. et al., 2015. Miniaturized mitogenome of the parasitic plant *Viscum scurruloideum* is extremely divergent and dynamic and has lost all nad genes. *Proceedings of the National Academy of Sciences of the United States of America*, 112(27), pp.E3515-3524.

Skladal, D., Halliday, J. & Thorburn, D.R., 2003. Minimum birth prevalence of mitochondrial respiratory chain disorders in children. *Brain*, 126(8), pp.1905–1912.

Sloan, D.B. et al., 2012. Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. *PLoS Biology*, 10(1), pp. e1001241

Stehling, O., Wilbrecht, C. & Lill, R., 2014. Mitochondrial iron-sulfur protein biogenesis and human disease. *Biochimie*, 100, pp.61–77.

Stroeher, E. et al., 2015. Glutaredoxin S15 is involved in Fe-S cluster transfer in mitochondria influencing lipoic acid-dependent enzymes, plant growth and arsenic tolerance in *Arabidopsis*. *Plant Physiology*, 170(March), pp.1284–1299.

Ströher, E. & Millar, A. H., 2012. The biological roles of glutaredoxins. *Biochemical Journal*, 446(3), pp.333–348.

Stroud, D.A. et al., 2016. Accessory subunits are integral for assembly and function of human mitochondrial complex I. *Nature*, 538(7623), pp.123–126.

Subrahmanian, N., Remacle, C. & Hamel, P.P., 2016. Plant mitochondrial Complex I composition and assembly: a review. *Biochimica et Biophysica Acta - Bioenergetics*, 1857(7), pp.1001-1014

- Sugiana, C. et al., 2008. Mutation of C20orf7 Disrupts Complex I Assembly and Causes Lethal Neonatal Mitochondrial Disease. *American Journal of Human Genetics*, 83(4), pp.468–478.
- Sun, F. et al., 2015. Empty pericarp7 encodes a mitochondrial E-subgroup pentatricopeptide repeat protein that is required for ccmFN editing, mitochondrial function and seed development in maize. *Plant Journal*, 84(2), pp.283–295.
- Swalwell, H. et al., 2011. Respiratory chain complex I deficiency caused by mitochondrial DNA mutations. *European Journal of Human Genetics*, 19(7), pp.769–775.
- Sweetlove, L.J., Taylor, N.L. & Leaver, C.J., 2007. Isolation of intact, functional mitochondria from the model plant *Arabidopsis thaliana*. *Methods in Molecular Biology (Clifton, N.J.)*, 372(2), pp.125–136.
- Tenisch, E.V. et al., 2012. Massive and exclusive pontocerebellar damage in mitochondrial disease and NUBPL mutations. *Neurology*, 79(4), p.391.
- Tielens, A.G.M. et al., 2002. Mitochondria as we don't know them. 27(11), pp.564–572.
- Tocilescu, M.A. et al., 2010. The role of a conserved tyrosine in the 49-kDa subunit of complex I for ubiquinone binding and reduction. *Biochimica et Biophysica Acta - Bioenergetics*, 1797(6–7), pp.625–632.
- Tocilescu, M. a. et al., 2007. Exploring the ubiquinone binding cavity of respiratory complex. *Journal of Biological Chemistry*, 282, pp.29514–29520.
- Tonduti, D. et al., 2015. New spastic paraplegia phenotype associated to mutation of NFU1. *Orphanet Journal of Rare Diseases*, 10, p.1-3.
- Tort, F., Ferrer-Cortes, X. & Ribes, A., 2016. Differential diagnosis of lipoic acid synthesis defects. *Journal of Inherited Metabolic Disease*, pp.781–793.
- Touraine, B. et al., 2004. Nfu2: a scaffold protein required for [4Fe-4S] and ferredoxin iron-sulphur cluster assembly in *Arabidopsis* chloroplasts. *The Plant Journal*, 40(1), pp.101–111.
- Tovar, J., Fischer, A. & Clark, C.G., 1999. The mitosome, a novel organelle related to mitochondria in the amitochondrial parasite *Entamoeba histolytica*. *Molecular Microbiology*, 32, pp.1013–1021.



- Towbin, H., Staehelin, T. & Gordon, J., 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences of the United States of America*, 76(9), pp.4350–4354.
- Triepels, R.H. et al., 2001. Human Complex I Defects Can Be Resolved by Monoclonal Antibody Analysis into Distinct Subunit Assembly Patterns. *Journal of Biological Chemistry*, 276(12), pp.8892–8897.
- Tucker, E.J. et al., 2012. Next-generation sequencing in molecular diagnosis: NUBPL mutations highlight the challenges of variant detection and interpretation. *Human Mutation*, 33(2), pp.411–418.
- Uzarska, M.A. et al., 2016. Mitochondrial Bol1 and Bol3 function as assembly factors for specific iron-sulfur proteins. *eLife*, 5, pp.e16673.
- Uzarska, M.A. et al., 2013. The mitochondrial Hsp70 chaperone Ssq1 facilitates Fe/S cluster transfer from Isu1 to Grx5 by complex formation. *Molecular Biology of the Cell*, 24(12), pp.1830–1841.
- Vafai, S.B. & Mootha, V.K., 2012. Mitochondrial disorders as windows into an ancient organelle. *Nature*, 491(7424), pp.374–383.
- Varadarajan, J. et al., 2010. ATR3 encodes a diflavin reductase essential for Arabidopsis embryo development. *New Phytologist*, 187(1), pp.67–82.
- Varghese, F. et al., 2015. Characterization of clinically identified mutations in NDUFV1, the flavin-binding subunit of respiratory complex I, using a yeast model system. *Human Molecular Genetics*, 24(22), pp.6350–6360.
- Vazzola, V. et al., 2007. Knockout of frataxin gene causes embryo lethality in Arabidopsis. *FEBS Letters*, 581(4), pp.667–672.
- Videira, A. & Duarte, M., 2002. From NADH to ubiquinone in Neurospora mitochondria. *Biochimica et Biophysica Acta - Bioenergetics*, 1555(1–3), pp.187–191.
- Vigani, G., Maffi, D. & Zocchi, G., 2009. Iron availability affects the function of mitochondria in cucumber roots. *New Phytologist*, 182(1), pp.127–136.
- Vinothkumar, K.R., Zhu, J. & Hirst, J., 2014. Architecture of mammalian respiratory complex I. *Nature*, 515(7525), pp.80–84.

Vogel, R.O. et al., 2005. Human mitochondrial complex I assembly is mediated by NDUFAF1. *FEBS Journal*, 272(20), pp.5317–5326.

Vogel, R.O., Smeitink, J.A.M. & Nijtmans, L.G.J., 2007. Human mitochondrial complex I assembly: A dynamic and versatile process. *Biochimica et Biophysica Acta - Bioenergetics*, 1767(10), pp.1215–1227.

Vogel, R.O., Janssen, R.J.R.J., et al., 2007b. Cytosolic signaling protein Ecsit also localizes to mitochondria where it interacts with chaperone NDUFAF1 and functions in complex I assembly. *Genes and Development*, 21(5), pp.615–624.

Wallace, D.C. et al., 1988. Mitochondrial DNA Mutation Associated with Leber's Hereditary Optic Neuropathy. *Science*, 242(4884), pp.1427–1430.

Waller, J.C. et al., 2012. Mitochondrial and plastidial COG0354 proteins have folate-dependent functions in iron-sulphur cluster metabolism. *Journal of Experimental Botany*, 63(1), pp.403–411.

Wessels, H.J.C.T. et al., 2013. Analysis of 953 Human Proteins from a Mitochondrial HEK293 Fraction by Complexome Profiling. *PLoS ONE*, 8(7), pp.e68340

Wessels, H.J.C.T. et al., 2009. LC-MS/MS as an alternative for SDS-PAGE in blue native analysis of protein complexes. *Proteomics*, 9(17), pp.4221–4228.

Westwood, J.H. et al., 2010. The evolution of parasitism in plants. *Trends in Plant Science*, 15(4), pp.227–235.

Wingert, R. A et al., 2005. Deficiency of glutaredoxin 5 reveals Fe-S clusters are required for vertebrate haem synthesis. *Nature*, 436(7053), pp.1035–1039.

Wirth, C. et al., 2016. Structure and function of mitochondrial complex I. *Biochimica et Biophysica Acta - Bioenergetics*, 1857(7), pp.902–914.

Wittig, I., Braun, H.-P. & Schagger, H., 2006. Blue native PAGE. *Nature Protocols*, 1, pp.418–428.

Wydro, M.M. et al., 2013. The evolutionarily conserved iron-sulfur protein INDH is required for complex I assembly and mitochondrial translation in Arabidopsis [corrected]. *Plant Cell*, 25(10), pp.4014–4027.

- Wydro, M.M. & Balk, J., 2013. Insights into the pathogenic character of a common NUBPL branch-site mutation associated with mitochondrial disease and complex I deficiency using a yeast model. *Disease Models & Mechanisms*, 6(5), pp.1279–1284.
- Xi, Z. et al., 2013. Massive Mitochondrial Gene Transfer in a Parasitic Flowering Plant Clade. *PLoS Genetics*, 9(2), pp.1–10.
- Xu, X.M. et al., 2009. Dual localized AtHscB involved in iron sulfur protein biogenesis in arabidopsis. *PLoS ONE*, 4(10), pp.e7662
- Xu, X.M. & Moller, S.G., 2004. AtNAP7 is a plastidic SufC-like ATP-binding cassette/ATPase essential for Arabidopsis embryogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, 101(24), pp.9143–9148.
- Xu, X.M. & Møller, S.G., 2006. AtSufE is an essential activator of plastidic and mitochondrial desulfurases in Arabidopsis. *The EMBO Journal*, 25(4), pp.900–909.
- Yip, C. et al., 2011. Evolution of Respiratory Complex I. *Journal of Biological Chemistry*, 286(7), pp.5023–5033.
- Yuvaniyama, P. et al., 2000. NifS-directed assembly of a transient [2Fe-2S] cluster within the NifU protein. *Proceedings of the National Academy of Sciences of the United States of America*, 97(2), pp.599–604.
- Zhu, J., Vinothkumar, K.R. & Hirst, J., 2016. Structure of mammalian respiratory complex I. *Nature*, 536(7616), pp.354–358.
- Zickermann, V., Wirth, C. & Nasiri, H., 2015. Mechanistic insight from the crystal structure of mitochondrial complex I. *Science*, 347(6217), pp.4–9.
- Zonneveld, B.J.M., 2010. New record holders for maximum genome size in eudicots and monocots. *Journal of Botany*, 2010(April).
- Zuber, D., 2004. Biological flora of Central Europe: *Viscum album* L. *Flora - Morphology, Distribution, Functional Ecology of Plants*, 199, pp.181–203.

## 8 Appendix

### Proteomic data from Chapter 4

Proteomic data, discussed in chapter 4, displayed as the output from the MASCOT server. For each protein, the total number of peptides and number of unique peptides that were significant ( $p < 0.05$ ) is shown. For table 8.3 and 8.4, partial lists are displayed, showing proteins involved in the respiratory process. For the dataset in table 8.3, there were 587 protein hits in total. For the dataset in table 8.4, there were 684 protein hits.

**Table 8.1: mtDNA encoded mitochondrial proteins from *Viscum album* Percoll-purified mitochondria**

Gene	Description	Total peptides (significant)	Unique peptides (significant)
atp1	atp1 ATP synthase F1 subunit 1	42	14
cox2	cox2 cytochrome c oxidase subunit 2	4	2
cob	cob apocytochrome b	5	2
atp8	atp8 ATP synthase F0 subunit 8	3	2
atp4	atp4 ATP synthase F0 subunit 4	2	2
cox3	cox3 cytochrome c oxidase subunit 3	1	1
rps3	rps3 ribosomal protein S3	1	1
cox1	cox1 cytochrome c oxidase subunit 1	1	1

**Table 8.2: mtDNA encoded mitochondrial proteins from *Viscum album* crude mitochondria**

Gene	Description	Total peptides (significant)	Unique peptides (significant)
atp1	atp1 ATP synthase F1 subunit 1	86	19
cox2	cox2 cytochrome c oxidase subunit 2	8	4
atp4	atp4 ATP synthase F0 subunit 4	6	4
cob	cob apocytochrome b	5	3
atp8	atp8 ATP synthase F0 subunit 8	3	2
rps3	rps3 ribosomal protein S3	4	4
cox1	cox1 cytochrome c oxidase subunit 1	3	2
cox3	cox3 cytochrome c oxidase subunit 3	3	2
rps4	rps4 ribosomal protein S4	4	3
rpl16	rpl16 ribosomal protein L16	1	1

**Table 8.3: Mitochondrial proteins identified based on the *Arabidopsis thaliana* database in *Viscum album* Percoll purified mitochondria**

Gene	Description	Total peptides (significant)	Unique peptides (significant)
AT5G08670.1	ATP synthase alpha/beta family protein	14	5
ATMG01190.2	edited ATP1	16	6
AT5G66760.1	SDH1-1 succinate dehydrogenase 1-1	9	6
AT4G21490.1	NDB3 NAD(P)H dehydrogenase B3	3	2
AT3G27240.1	Cytochrome C1 family	2	1
ATMG00160.1	COX2 cytochrome oxidase 2	1	1

**Table 8.4: Mitochondrial proteins identified based on the *Arabidopsis thaliana* database in *Viscum album* crude mitochondria**

Gene	Description	Total peptides (significant)	Unique peptides (significant)
AT5G08670.1	ATP synthase alpha/beta family protein	39	8
ATMG01190.2	edited ATP1	20	7
AT5G66760.1	SDH1-1 succinate dehydrogenase 1-1	15	7
AT5G64210.1	AOX2 alternative oxidase 2	2	1
AT5G57815.1	Cytochrome c oxidase, subunit Vib family protein	5	2
AT5G40810.1	Cytochrome C1 family	5	2
AT4G21490.1	NDB3 NAD(P)H dehydrogenase B3	4	3
ATMG00160.1	COX2 cytochrome oxidase 2	1	1
AT5G13430.1	Ubiquinol-cytochrome C reductase iron-sulfur subunit	3	3
AT4G10040.1	CYTC-2 cytochrome c-2	1	1
ATMG01170.1	ATP6-2 ATPase, F0 complex, subunit A protein	1	1
AT5G08740.1	NDC1 NAD(P)H dehydrogenase C1	2	2
AT5G40650.1	SDH2-2 succinate dehydrogenase 2-2	1	1
AT3G22370.1	AOX1A, ATAOX1A alternative oxidase 1A	1	1